12-1-05

IPW

Express Mail Label No. EV695506153US
Date of Deposit: November 29, 2005

Attorney Docket No. 25371-021CIP

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT:	Nakamura, et al.		
SERIAL NUMBER:	10/788,847	EXAMINER :	Not Yet Assigned
FILING DATE:	February 27, 2004	ART UNIT :	1645
FOR:	GENE AND PROTEIN RELATING TO HEPATOCELLULAR CARCINOMA AND METHODS OF USE THEREOF		

Commissioner for Patents
P. O. Box 1450
Alexandria, VA 22313-1450

TRANSMITTAL

Transmitted herewith for filing in the above-referenced patent application are the following documents:

1. Communication (1 pg);
2. Certified copy of Canadian Patent Application No. 2,399,569 (50 pgs); and
3 Return Postcard

Although Applicants believe that no additional fees are due in connection with this submission, the Commissioner is authorized to charge any deficiencies to Deposit Account No. 50-0311 (Reference No. 25371-021CIP.

If the enclosed papers are considered incomplete, the Mail Room is respectfully requested to contact the undersigned collect at (617) 542-6000, Boston, Massachusetts.

Respectfully submitted,

Ingrid A. Beattie, Reg. No. 42,306
Cynthia A. Kozakiewicz, Reg. No. 42,764
c/o MINTZ LEVIN
Tel: (617) 542-6000
Fax: (617) 542-2241
Customer No. 30623

Dated: November 29, 2005

Express Mail Label No. EV695506153US
Date of Deposit: November 29, 2005

Attorney Docket No. 25371-021CIP

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT:	Nakamura, et al.		
SERIAL NUMBER:	10/788,847	EXAMINER :	Not Yet Assigned
FILING DATE:	February 27, 2004	ART UNIT :	1645
FOR:	GENE AND PROTEIN RELATING TO HEPATOCELLULAR CARCINOMA AND METHODS OF USE THEREOF		

Commissioner for Patents
P. O. Box 1450
Alexandria, VA 22313-1450

COMMUNICATION

Enclosed is a certified copy of Canadian Patent Application 2,399,569, the priority document of the above-referenced application.

The Commissioner is authorized to charge any fees that may be due, or to credit any overpayment, to the undersigned's account, Deposit Account No. 50-0311 Ref. No. 25371-021 CIP.

Respectfully submitted,

Ingrid A. Beattie, Reg. No. 42,306
Cynthia A. Kozakiewicz, Reg. No. 42,764
c/o MINTZ LEVIN
Tel: (617) 542-6000
Fax: (617) 542-2241
Customer No. 30623

Dated: November 29, 2005

Office de la propriété intellectuelle du Canada

Un organisme d'Industrie Canada

Canadian Intellectual Property Office

An Agency of Industry Canada

Bureau canadien des brevets

Certification

Canadian Patent Office

Certification

La présente atteste que les documents ci-joints, dont la liste figure ci-dessous, sont des copies authentiques des documents déposés au Bureau des brevets.

This is to certify that the documents attached hereto and identified below are true copies of the documents on file in the Patent Office.

Specification and Drawings, as originally filed, with Application for Patent Serial No: **2,399,569,** on August 23, 2002, by **YUSUKE NAKAMURA and YOICHI FURUKAWA,** for "Diagnostic Markers and Drug Targets for Treatment of Cancer".

Sylvie Grégoire

Agent certificateur/Certifying Officer

November 23, 2005

Date

Canada

(CIPO 68)
31-03-04

OPIC CIPO

Abstract

To disclose mechanisms of hepatocellular carcinogenesis and to identify novel diagnostic markers and/or drug targets for treatment of hepatocellular carcinomas (HCCs), we analyzed expression profiles of clinical HCCs using a genome-wide cDNA microarray. From among the transcripts that were commonly up-regulated in these tumors, we identified a novel human gene, at chromosomal band 1p36.13, termed *DDEFL1* (development and differentiation enhancing factor-like 1) encoding a product that shared structural features with centaurin-family proteins as well as *VANGL1*, *LGN*, and a novel gene termed *ZNFN3A1* at chromosomal band 1q44. The deduced 903-amino-acid sequence of DDFEL1 showed 46% homology to DDEF/ASAP1 (development and differentiation enhancing factor), and contained an Arf GTPase-activating protein (ArfGAP) domain and two ankyrin repeats. The predicted 428-amino-acid sequence of ZNFN3A1 contained a zf-MYND (MYND domain containing zinc finger protein) domain and a SET (Su (var) 3-9, Enhancer-of-zeste, Trithorax) domain. *VANGL1* (Van Gogh Like 1) is a gene homologous to *strabismus* (*Van Gogh*) that is involved in cell polarity and cell fate decisions in *Drosophila*. LGN protein interacts with alpha subunit of inhibitory heterotrimeric G proteinis ($G\alpha_{i2}$). Gene transfer of *DDEFL1*, *ZNFN3A1* or *LGN* promoted proliferation of cells that lacked endogenous expression of either of these genes. Furthermore, reduction of *DDEFL1*, *ZNFN3A1*, *VANGL1* or *LGN* expression by transfection of their specific anti-sense S-oligonucleotides inhibited the growth of hepatocellular carcinoma cells. Our results provide novel insight into hepatocarcinogenesis and may contribute to development of new strategies for diagnosis and treatment of HCC.

DIAGNOSTIC MARKERS AND DRUG TARGETS FOR TREATMENT OF CANCER

Introduction

Hepatocellular carcinoma (HCC) is a leading cause of cancer deaths worldwide. In spite of recent progress in therapeutic strategies, prognosis of patients with advanced HCC remains very poor. Although molecular studies have revealed that alterations of *TP53*, *CTNNB1* and/or *AXIN1* genes can be involved in hepatocarcinogenesis (1, 2), these changes appear to be implicated in only a fraction of HCCs. To disclose mechanisms of hepatocellular carcinogenesis from a genome-wide point of view and to discover target molecules for diagnosis and for development of novel therapeutic drugs, we have been analyzing expression profiles of HCCs by means of a cDNA microarray representing 23,040 genes. These efforts have pinpointed 165 genes, including 69 ESTs, which appear to be up-regulated frequently in cancer tissues compared with corresponding non-cancerous liver cells. Since carcinogenesis involves activation of oncogenes and/or inactivation of tumor suppressor genes, enhanced expression of at least some of these 165 genes may reflect oncogenic properties.

In this paper we describe isolation of two novel human genes, *DDEFL1* (development and differentiation enhancing factor-like 1) and *ZNFN3A1* (Zinc Finger protein N3A1), and two human genes, *VANGL1* and *LGN* from among the transcripts whose expression was frequently elevated in HCCs. Reducing the expression of *DDEFL1*, *ZNFN3A1*, *VANGL1* or *LGN* by transfection of anti-sense DNA was able to inhibit growth of human HCC cells *in vitro*. These data should shed light on the mechanisms of hepatocellular carcinogenesis and also provide clues for development of novel diagnostic and therapeutic strategies.

Materials and Methods

Patients and tissue specimens

All HCC tissues and corresponding non-cancerous tissues were obtained with informed consent from surgical specimens of patients who underwent hepatectomy in Kyoto University Hospital.

Cell lines

Human embryonic kidney 293 (HEK293) cells, COS7 cells, NIH3T3 cells and human hepatoma cell lines HepG2, Huh7 and Alexander were obtained from the American Type Culture Collection (ATCC). SNU423, SNU449, SNU475 were obtained from the Korea cell-line bank. All cells were grown in monolayers in appropriate media: Dulbecco's modified Eagle's medium for COS7, NIH3T3, HEK293, HepG2, Huh7 and Alexander and RPMI1640 for SNU423, SNU449 and SNU475; both media were supplemented with 10% fetal bovine serum and 1% antibiotic/antimycotic solution (Sigma).

RNA preparation and RT-PCR

Total RNA was extracted with a Qiagen RNeasy kit (Qiagen) or Trizol reagent (Life Technologies, Inc.) according to the manufacturers' protocols. Ten-microgram aliquots of total RNA were reversely transcribed for single-stranded cDNAs using poly dT_{12-18} primer (Amersham Pharmacia Biotech) with Superscript II reverse transcriptase (Life Technologies). Each single-stranded cDNA preparation was diluted for subsequent PCR amplification by standard RT-PCR experiments carried out in 20-μl volumes of PCR buffer (TAKARA). Amplification proceeded for 4 min at 94°C for denaturing, followed by 20 (for *GAPDH*) or 33 (for *DDEFL1*) cycles of 94°C for 30 s, 56°C for 30 s, and 72°C for 45 s, in the GeneAmp PCR system 9700 (Perkin-Elmer, Foster City, CA). Primer sequences were; for *GAPDH:* forward, 5'-ACAACAGCCTCAAGA TCATCAG and reverse, 5'-GGTCCACCACTGACACGTTG; for *DDEFL1*: forward, 5'-AGCTGAGACATTTGTTCTCTTG and reverse: 5'-TATAAACCAG CTGAGTCCAGAG; for *ZNFN3A1* forward: 5'-TTCCCGATATCAACATC TACCAG reverse: 5'-AGTGTGTGACCTCAATAAGGCAT; for *VANGL1* forward: 5'-GAGTTGTATTATGAAGAGGCCGA reverse: 5'- ATGTCTCAGACTGTAA

GCGAAGG; and for *LGN* forward: 5'-ATCTGAAGCACTTAGCAATTGC reverse: 5'-CTGTAGCTCAGACCAAGAACC.

Northern-blot analysis

Human multiple-tissue blots (Clontech, Palo Alto, CA) were hybridized with a ^{32}P-labeled cDNA designated *DDEFL1*, *ZNFN3A1* and *VANGL1*. Pre-hybridization, hybridization and washing were performed according to the supplier's recommendations. The blots were autoradiographed with intensifying screens at –80°C for 72 h.

5' rapid amplification of cDNA ends (5' RACE)

5' RACE experiments were carried out using a Marathon cDNA amplification kit (Clontech, Palo Alto, CA) according to the manufacturer's instructions. For the amplification of the 5' part of *DDEFL1*, *ZNFN3A1*, and *VANGL1* cDNAs, gene-specific reverse primers (5'-CTCACTTGGCACGTCAG CAGGG, 5'-CTGCCAAGAAGTCGGAGTCTGGAG and 5'-TGTCAGCTCTCCG CTTGCGGAAAAAAAG, respectively) and the AP-1 primer supplied in the kit were used. The cDNA template was synthesized from human liver mRNA. The PCR products were cloned using a TA cloning kit (Invitrogen) and their sequences were determined with an ABI PRISM 3700 DNA sequencer (Applied Biosystems).

Establishment of cells over-expressing *DDEFL*, *ZNFN3A1* or *LGN*

The entire-coding sequences of *DDEFL1*, *ZNFN3A1*, *LGN* or *LacZ* were cloned separately into the pcDNA3.1-myc/His, or pcDNA 3.1 vector (Invitrogen),. Each clone was transfected into NIH3T3 and SNU423 cells using FuGENE6 reagent (Boehringer) according to the supplier's recommendations. 24 h after transfection, geneticin was added to the cultures and single colonies were selected two weeks after transfection. Expression of *DDEFL1*, *ZNFN3A1*, and *LGN* was determined by semi-quantitative RT-PCR.

Immunoblotting

Cells transfected with pcDNA3.1-myc/His-DDEFL1, pEGFP-ZNFN3A1, pFLAG-ZNFN3A1, pcDNA3.1-myc/His-VANGL1 or pcDNA3.1-myc/His-LGN were washed

twice with PBS and harvested in lysis buffer (150 mM NaCl, 1% Triton X-100, 50 mM Tris-HCl pH 7.4, 1mM DTT, and 1X complete Protease Inhibitor Cocktail (Boehringer)). After the cells were homogenized and centrifuged at 10,000x*g* for 30 min, the supernatant were standardized for protein concentration by the Bradford assay (Bio-Rad). Proteins were separated by 10% SDS-PAGE and immunoblotted with mouse anti-myc, anti-Flag, anti-GFP or anti-HA antibody. HRP-conjugated goat anti-mouse IgG (Amersham) served as the secondary antibody for the ECL Detection System (Amersham).

Immunohistochemical staining

Cells transfected with pcDNA3.1-myc/His-DDEFL1, pFLAG- ZNFN3A1 pcDNA3.1-myc/His-VANGL1 or pcDNA3.1-myc/His-LGN were fixed with PBS containing 4% paraformaldehyde for 15 min, then rendered permeable with PBS containing 0.1% Triton X-100 for 2.5 min at RT. Subsequently the cells were covered with 2% BSA in PBS for 24 h at 4°C to block non-specific hybridization. Mouse anti-myc monoclonal antibody (Sigma) at 1:1000 dilution or mouse anti-FLAG antibody (Sigma) at 1:2000 dilution was used for the first antibody, and the reaction was visualized after incubation with Rhodamine-conjugated anti-mouse second antibody (Leinco and ICN). Nuclei were counter-stained with 4',6'-diamidine-2'-phenylindole dihydrochloride (DAPI). Fluorescent images were obtained under an ECLIPSE E800 microscope.

Colony-formation assay

Cells plated onto 10-cm dishes ($2X10^5$ cells/dish) were transfected either with plasmid or with synthetic S-oligonucleotides, and cultured with an appropriate concentration of geneticin for two weeks. The cells were then fixed with 100% methanol and stained by Giemsa solution. Sense or antisense S-oligonucleotides of *DDEFL1*, encompassing the transcription-initiation site or the first ATG sequence, were transfected using LIPOFECTIN Reagent (GIBCO BRL). Sequences of the S-oligonucleotides were as follows: *DDEFL1* sense, 5'-CCTC GCGTGGTGAGTA; *DDEFL1* antisense, 5'-TACTCACCACGCGAGG, *ZNFN3A1* sense; 5'-GGCTCCATCCTCCCGC, *ZNFN3A1* antisense; 5'-GCGGGAGGATG GAGCC, *VANGL1* sense; 5'-ATACCCAATCCACTTA, *VANGL1* antisense; 5'-

TAAGTGGATTGGGTAT, *LGN* sense; 5'-CGATGGAGGAAAATTT, and *L G N* antisense; 5'-AAATTTTCCTCCATCG .

Flow cytometry

Cells were plated at a density of $1X10^5$ cells/100 mm dish and trypsinized at the given time course, followed by fixation in 70% cold ethanol. After RNase treatment, cells were stained with propidium iodide (50 µg/ml) in PBS. Flow cytometry was performed on a Becton Dickinson FACScan and analyzed by CellQuest and ModFit software (Verity Software House). The percentages of nuclei in G0/G1, S and G2/M phases of the cell cycle, and any sub-G1 population were determined from at least 20,000 ungated cells.

3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay

Cells were plated at a density of $5X10^5$ cells/100 mm dish. At 24 hours after seeding, the cells were transfected in triplicate with sense or antisense S-oligonucleotide designated to suppress either *ZNFN3A1 or VANGL1*. At 72 hours after transfection , the medium was replaced with fresh medium containing 500 µg/ml of M T T (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide) (Sigma) and the plates were incubated for four hours at 37°C. Subsequently, the cells were lysed by the addition of 1 ml of 0.01 N HCl/10%SDS and. absorbance of lysates was measured with an ELISA plate reader at a test wavelength of 570 nm (reference, 630 nm). The cell viability was represented by the absorbance compared to that of control cells.

Results

Identification of *DDEFL1*, a commonly up-regulated gene in human hepatocellular carcinomas

By means of a genome-wide cDNA microarray containing 23040 genes, we compared expression profiles of 20 HCCs with their corresponding non-cancerous liver tissues. A gene with an in-house accession number of B9362 corresponding to an EST, Hs.44579 of a UniGene cluster (http://www.ncbi. nlm.nih.gov/UniGene/), was found to be over-expressed in a range between 1.57 and 5.83 (Fig. 1a). Since an open reading frame of this gene encoded a protein approximately 60% identical

to that of DDEF2_(development and differentiation enhancing factor_2), we termed this gene *DDEFL1* (development and differentiation enhancing factor-like 1). To clarify the results of the cDNA microarray, we examined expression of this transcript in an additional 11 HCCs by semi-quantitative RT-PCR and confirmed its increased expression in nine of these tumors (Fig. 1b).

Isolation and structure of a novel gene *DDEFL1*

Multiple-tissue northern-blot analysis using a PCR product of *DDEFL1* as a probe revealed a 4-kb transcript that was expressed in lung, liver, small intestine, placenta and peripheral blood leukocyte (Fig. 2a). Since *B9362* was smaller than that detected on the Northern blot, we carried out 5'RACE experiments to determine the entire coding sequence of the gene. The complete cDNA consisted of 4050 nucleotides, with an open reading frame of 2712 nucleotides encoding a 903-amino-acid protein (GenBank accession number AB051853). The first ATG was flanked by a sequence (CCCGCCATGC) that agreed with the consensus sequence for initiation of translation in eukaryotes, with an in-frame stop codon upstream. Using the BLAST program to search for homologies in the NCBI (the National Center for Biotechnology Information) database, we identified a genomic sequence with GenBank accession number AL357134, which had been assigned to chromosomal band 1p36.12. Comparison of the cDNA and genomic sequences disclosed that the novel gene consisted of 25 exons (Fig. 2b).

A search for protein motifs with the Simple Modular Architecture Research Tool (SMART, http://smart.embl-heidelberg.de) revealed that the predicted protein contained two coiled-coil regions (codons 141-172 and 241-278), a PH (Pleckstrin homology) motif (codons 303-396), a motif of ArfGAP (GTPase-activating protein for Arf) (codons 426-551) and two ankyrin repeats (codons 585-617 and 621-653). This structure was similar to centaurin beta 1 and centaurin beta 2 (Fig. 2c). In particular, DDEFL1 shared features of centaurin-family proteins such as a PH domain, a target of phosphatidylinositol 3,4,5-trisphosphate, and a motif of ArfGAP. The amino acid sequence of the ArfGAP motif of DDEFL1 was 67.8% identical to that of DDEF2 (Fig. 2d). Notably, the $CXXCX_{16}CXXC$ motif, representing a zinc finger structure essential to GAP activity, was completely preserved.

Subcellular localization of DDEFL1

We transiently transfected a plasmid expressing myc-tagged DDEFL1 protein (pDNA-myc/His-DDEFL1) into COS7 cells and subjected the cells to immunocytochemical staining. The expected myc-tagged protein was detected on western blots with an anti-myc antibody (Fig. 3a), and subsequent microscopic analysis indicated that the protein was present mainly in the cytoplasm (Fig. 3b). DDEFL1 was also localized in the cytoplasm of human embryonal kidney (HEK293) cells.

Effect of *DDEFL1* on cell growth

We carried out a colony-formation assay by transfecting NIH3T3 cells with a plasmid expressing *DDEFL1* (pcDNA-DDEFL1). Cells transfected with pcDNA-DDEFL1 produced markedly more colonies than control cells. An increase in colony formation similarly occurred with transfected human hepatoma SNU423 and Alexander cells, in which endogenous expression of *DDEFL1* is very low (Fig. 4a).

To investigate this growth-promoting effect further, we established NIH3T3 cells that stably expressed exogeneous *DDEFL1* (Fig.4b). The growth rate of NIH3T3-DDEFL1 cells was significantly higher than that of parental NIH3T3 cells in culture media containing 10% FBS (Fig. 4c); but although NIH3T3-DDRFL1 cells survived for 6 days in media containing only 0.1% FBS, control NIH3T3 cells died within 6 days under the same conditions (Fig. 4d).

Suppression of *DDEFL1* expression in human hepatoma SNU475 cells by antisense S-oligonucleotides

Six pairs of control and antisense S-oligonucleotides corresponding to the *DDEFL1* gene were transfected into SNU475 cells, which had shown the highest level of *DDEFL1* expression among the six hepatoma cell lines we examined. Antisense S-oligonucleotides #1 and #5 significantly suppressed expression of *DDEFL1* compared to control S-oligonucleotides, 12 and 24 hours after transfection (Fig. 5a). Six days after transfection, surviving cells transfected with antisense S-oligonucleotides #1 or #5 were markedly fewer than cells transfected with control S-oligonucleotides (Fig. 5b). Consistent results were obtained in three independent experiments.

Identification of *A6681*, a novel gene frequently up-regulated in HCCs

Among commonly up-regulated genes by the microarray analysis, we also focused on a gene, A6681, corresponding to an EST (Hs.8109) of a UniGene cluster (http://www.ncbi.nlm.nih.gov/UniGene/), because its expression was significantly up-regulated in eleven of twelve clinical HCCs compared with the corresponding non-cancerous liver tissues (Fig. 6a). The elevated expression of A6681 was also confirmed in additional ten HCC cases by semi-quantitative RT-PCR (Fig. 6b).

Expression, isolation and characterization of a novel human gene, *ZNFN3A1*

Multi-tissue northern blot analysis using A6681 cDNA as a probe showed a 1.7-kb transcript expressed specifically in testis and skeletal muscle (Fig. 7a). We searched for genomic sequences corresponding to A6681 in the genomic databases and found a draft sequence assigned to chromosomal band 1q44 (GenBank accession number:NT004536). Using GENSCAN and Gene Recognition and Assembly Internet Link program, we obtained candidate-exon sequences and performed exon-connection using these sequence information. In addition, we carried out 5'-RACE to determine the sequence of the 5' region of the transcript. As a result, we obtained an assembled sequence of 1622 nucleotides containing an open reading frame of 1284 nucleotides (GenBank accession number:AB057595). Simple Modular Architecture Research Tool suggested that the predicted protein contained a zf-MYND [zinc finger protein (MYND domain containing)] domain (codons 49-87) and a SET [(Su (var) 3-9, Enhancer-of-zeste, Trithorax)] domain (codons 117-246) (Fig. 7b). The gene was termed *ZNFN3A1* (zinc finger protein, subfamily 3A (MYND domain containing), 1) by the nomenclature committee.

Subcellular localization of ZNFN3A1

The entire coding region corresponding to *ZNFN3A1* was cloned into a pEGFP-N1 vector and a pFlag-CMV-5a vector, and these constructs were transiently transfected into SNU475 cells. ZNFN3A1-EGFP fusion protein was present in the cytoplasm and nucleus by fluorescent microscopy. Similarly flag-tagged ZNFN3A1

protein was detected in the cytoplasm and nucleus by fluorescent immunohistochemical staining (Fig. 8).

Effect of *ZNFN3A1* on cell growth

To analyze the effect of *ZNFN3A1* on cell growth, we carried out a colony-formation assay by transfecting NIH3T3 cells with a plasmid expressing *ZNFN3A1* (pcDNA-ZNFN3A1). Compared with mock and antisense-*ZNFN3A1* expressing plasmids, sense-*ZNFN3A1* expressing plasmid induced markedly more colonies in NIH3T3 cells (Fig. 9a, b). This result was confirmed by three independent experiments. To investigate this growth-promotive effect of *ZNFN3A1*, we established NIH3T3 cells that stably expressed exogeneous *ZNFN3A1* (NIH3T3-ZNFN3A1 cells) (Fig. 9c). As shown Fig. 9d, growth rate of NIH3T3-ZNFN3A1 cells was higher than those transfected with antisense *ZNFN3A1* or mock cells.

Growth suppression of hepatoma cells by antisense S-oligonucleotides designated to reduce expression of *ZNFN3A1*

To test whether suppression of *ZNFN3A1* may result in cell cycle arrest and/or cell death of HCC cells, we synthesized various antisense S-oligonucleotides designated to suppress its expression. Among them, antisense S-oligonucleotides (AS) encompassing the initiation codon, but not other antisense or control S-oligonucleotides (SE) significantly decreased endogenous expression of *ZNFN3A1* in SNU475 and Huh7 cells that constitutively express *ZNFN3A1* abundantly (Fig. 10a). Transfection of the antisense S-oligonucleotides (AS) significantly reduced number of surviving cells compared with that of control sense S-oligonucleotides (SE) (Fig. 10b, c). This result was confirmed by three independent experiments. Furthermore, FACS analysis demonstrated that inhibition of *ZNFN3A1* significantly decreased number of cells in S phase and increased number in sub-G1 phase (Fig. 10d). These results suggest that *ZNFN3A1* may play an important role for cell growth and/or survival of hepatocellular carcinoma cells.

***D3244* is commonly increased in human hepatocellular carcinomas**

Additionally, we focused on a gene, *D3244* corresponding to an EST

(Hs.122730) of a UniGene cluster (http://www.ncbi.nlm.nih.gov/UniGene/), because it was also significantly up-regulated in ten of twelve clinical HCCs compared with the corresponding non-cancerous liver tissues. The relative expression ratio compared to corresponding non-cancerous tissue of these 12 tumors ranged from 1.5 to 16.0 (Fig. 11a). The elevated expression of *D3244* was also confirmed in ten additional HCC cases by semi-quantitative RT-PCR (Fig. 11b).

Expression of *VANGL1* in human adult tissues

Multi-tissue northern blot analysis using *D3244* cDNA as a probe showed a 1.9-kb transcript abundantly expressed in testis and ovary in a tissue-specific manner (Fig. 12a). We searched for genomic sequences corresponding to *D3244* in databases in NCBI (http://www.ncbi.nlm.nih.gov/), and found two sequences (GenBank accession number:AL450389 and AL592436) assigned to chromosomal band 1p22. Using GENSCAN, and Gene Recognition and Assembly Internet Link program, we predicted candidate-exon sequences and performed exon-connection. In addition, we carried out 5' RACE to determine the sequence of the 5' region of the transcript. As a result, we obtained an assembled human cDNA sequence of 1879 nucleotides containing an open reading frame of 1572 nucleotides (GenBank accession number: AB057596). The predicted amino acid sequence shared 40% and 63% identity with *strabismus* (*Van Gogh*) and *VANGL2*. Hence, we termed the gene corresponding *D3244* as *VANGL1* (Van Gogh Like 1). Simple Modular Architecture Research Tool suggested that the predicted protein contained putative four transmembrane domains (codons 111-133,148-170,182-204, 219-241) (Fig. 12b).

Subcellular localization of VANGL1

The pcDNA3.1-myc/His-VANGL1 plasmid expressing c-myc-tagged VANGL1 protein was transiently transfected into SNU475 cells. Immunocytochemical staining revealed that the taggedVANGL1 protein was present in the cytoplasm (Fig. 13).

Growth suppression of hepatoma cells by antisense S-oligonucleotides designated to reduce expression of *VANGL1*

To test whether suppression of *VANGL1* may result in cell cycle arrest and/or cell death of HCC cells, we synthesized antisense S-oligonucleotides to suppress its expression. Among them, antisense S-oligonucleotides encompassing the initiation codon significantly decreased endogenous expression of *VANGL1* in SNU475 cells (Fig. 14a). Transfection of the antisense S-oligonucleotides significantly reduced number of surviving cells compared with control sense S-oligonucleotides (Fig. 14b, c). This result was confirmed by three independent experiments. Furthermore, FACS analysis demonstrated that inhibition of *VANGL1* significantly increased number of cells at sub-G1 phase (Fig. 14d). These results suggest that *VANGL1* may play an important role for cell growth and/or survival of hepatocellular carcinoma cells.

LGN is commonly increased in human hepatocellular carcinomas

Among commonly up-regulated genes by the microarray analysis, we further selected a gene, *D3636* corresponding to *LGN* (GenBank accession number: U54999), because it was significantly up-regulated in ten of twelve clinical HCCs compared with the corresponding non-cancerous liver tissues. The relative expression ratio compared to corresponding non-cancerous tissue of these 12 tumors ranged from 0.7 to 16.0 (Fig. 15a). The elevated expression of *LGN* was also confirmed in additional ten HCC cases by semi-quantitative RT-PCR (Fig. 15b).

Genomic structure of _LGN_

LGN cDNA consists of 2,336 nucleotides and encodes a 677 amino acid peptide. Comparison of the cDNA sequence with genomic sequences disclosed that the *LGN* gene consists of 14 exons (Fig. 16).

Subcellular localization of LGN

The pcDNA3.1-myc/His-LGN plasmid expressing c-myc-tagged LGN protein was transiently transfected into COS7 cells. A 72 kDa-band corresponding to myc-tagged LGN protein was detected by immunoblot analysis (Fig. 17). Immunocytochemical staining revealed that the tagged LGN protein was present in the cytoplasm and nucleus in the cells.

LGN gene transfer can promote cell growth

To analyze the effect of *LGN* on cell growth, we carried out a colony-formation assay by transfecting NIH3T3, SNU423, Alexander and SNU475 cells with a plasmid expressing *LGN* (pcDNA3.1-myc/His-LGN). Compared with a control plasmid (pcDNA3.1-myc/His-LacZ), pcDNA3.1-myc/His-LGN produced markedly a larger number of colonies in these cells (Fig. 18a). This result was confirmed by three independent experiments.

To further investigate the effect of *LGN* on cell growth, we established NIH3T3 cells that stably expressed exogeneous *LGN* (NIH3T3-*LGN* cells). NIH3T3-*LGN* cells showed higher growth rate than control NIH3T3-*LacZ* cells (Fig. 18b).

Antisense S-oligonucleotides of _LGN_ suppressed growth of human hepatoma SNU475 cells

We synthesized five pairs of sense and antisense S-oligonucleotides corresponding to *LGN* and then transfected into SNU423 cells. When we transfected antisense S-oligonucleotides, the antisense S-oligonucleotides #4 significantly suppressed expression of *LGN* compared to control S-oligonucleotides, 12 hours after transfection (Fig. 19a). Six days after transfection, the number of surviving cells transfected with antisense S-oligonucleotides #4 were markedly fewer than that with control S-oligonucleotides (Fig. 19b). Consistent results were obtained in three independent experiments.

Discussion

cDNA microarray technologies have enabled us to obtain comprehensive profiles of gene expression in normal versus malignant cells (1. 3-4). This approach discloses the complex nature of cancer cells, and helps to improve our understanding of carcinogenesis. Identification of genes that are deregulated in tumors can lead to more precise and accurate diagnosis of individual cancers, and to development of novel therapeutic targets (7). The genome-wide analysis of HCCs we reported previously disclosed a number of genes associated with

carcinogenesis, as well as genes affecting differentiation and vessel invasion. To identify ideal therapeutic targets for treatment of HCCs, we focused in the present work on *DDEFL1*, one of the genes up-regulated most frequently in HCCs according to our microarray analysis.

Since DDEFL1 showed 46% identity to a member of the centaurin family, DDEF2 (development and differentiation enhancing factor 2), a protein that regulates re-organization of the actin cytoskeleton, DDEFL1 may also play a role in organization of cellular structure (8). Because DDEFL1 also conserves a PH domain and an ArfGAP motif it appears to be a new member of the centaurin family, regulating Arf small GTPase by means of GAP activity. The PH domain, observed in the majority of molecules belonging to the Dbl family of GEFs, is thought to play a crucial role in relocation of proteins by interacting with specific target molecules and/or by directly regulating catalytic domains (5.9). Although DDEF2 is localized in peripheral focal adhesions, we found myc-tagged DDEFL1 protein to be diffuse in cytoplasm. Hence DDEFL1 may have a function different from DDEF2, but further investigation using specific antibody will be required to disclose the sub-cellular location(s) of DDEFL1 more precisely.

Arf proteins have been implicated in important cellular processes such as vesicular membrane transport, maintenance of the integrity of ER and Golgi compartments, and regulation of the peripheral cytoskeleton (6). Six members of Arf family (Arf1-Arf6) and their functions have been identified so far (11). For example, Arf6 proteins have been implicated as regulators of the cytoskeleton to alter the morphology of focal adhesions and to block spreading of cells, and DDEF2 displays GAP activity toward Arf1. To disclose the function of DDEFL1, however, it will be necessary to identify the molecular targets of its ArfGAP domain.

Since in our experiments over-expression of *DDEFL1* promoted growth promotion and survival of cells under low-serum conditions, *DDEFL1* may provide a growth advantage to cancer cells in poor nutritional and hypoxic conditions. Although the mechanism of this mitogenic effect remains to be investigated, the frequent up-regulation of *DDEFL1* in HCCs underscores the importance of this gene in hepatocarcinogenesis.

ZNFN3A1 has a zf-MYND domain and a SET domain, both of which are conserved in a putative proto-oncogene, *MTG8/ETO*. *MTG8/ETO* was identified by

molecular characterization of a t(8;21)(q22;q22) translocation breakpoint commonly found in the subset M2 of acute myeloid leukemias (12). The chimeric protein AML1/MTG8 produced by the translocation may play a crucial role in leukemogenesis by inactivating the function of AML1 and/or activating transcription through the zf-MYND domain of MTG8 (13, 14). Therefore up-regulation of *ZNFN3A1* may promote cell proliferation by regulating transcription of various genes as a component of transcriptional complex.

Strabismus (*stbm*) was identified as a gene responsible for a mutant fruit fly with rough eye phenotype (15). The gene is required to maintain polarity in the eye, legs and bristles and to decide cell fate of R3 and R4 photoreceptors in the *Drosophila*. A mouse gene homologous to *stbm*, *Ltap*, was altered in the neural tube mutant mouse Loop-tail, which is a human model of neural tube defects (NTDs) (16). Hence, VANGL1 may also play important roles in cellular polarity, cell fate decision, and/or organization of tissues. Since *VANGL1* is frequently up-regulated in HCCs and suppression of its expression significantly reduced growth or survival of cancer cells, VANGL1 may confer prolonged survival and/or depolarized growth to cancer cells.

LGN protein was previously reported as a protein interacting with alpha subunit of inhibitory heterotrimeric G proteins ($G\alpha_{i2}$) (17). The activating mutations of $G\alpha_{i2}$ have ever been reported in pituitary tumor and other endocrine tumors (18, 19, 20). However, involvement of *LGN* in tumorigenesis or carcinogenesis has not yet been reported. In this study, we showed that LGN was frequently up-regulated in HCCs compared with non-cancerous liver tissues. In addition, colony formation assay suggested that *LGN* might have oncogenic activity. Taken together, enhanced expression of *LGN* may activate $G\alpha_{i2}$ and mediate oncogenic signals in hepatocarcinogenesis. Further analysis of function of LGN and signal transduction pathway of heterotrimeric G proteins ($G\alpha_{i2}$) is required for a more profound understanding of hepatocarcinogenesis

We have demonstrated here that suppressing the expression of *DDEFL1*, *ZNFN3A1*, *VANGL1* or *LGN* by antisense oligonucleotides markedly decreases growth of HCC cells. Although the precise molecular mechanism by which the antisense S-oligonucleotides can suppress growth needs to be clarified, our data clearly indicate that these genes could be good candidates as diagnostic markers

for liver cancers, and could represent molecular targets for development of effective drugs to treat this often intractable disease.

References

1. Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., Fluge, O., Pergamenschikov, A., Williams, C., Zhu, S. X., Lonning, P. E., Borresen-Dale, A. L., Brown, P. O., and Botstein, D. Molecular portraits of human breast tumours, Nature. *406:* 747-752, 2000.

2. Satoh, S., Daigo, Y., Furukawa, Y., Kato, T., Miwa, N., Nishiwaki, T., Kawasoe, T., Ishiguro, H., Fujita, M., Tokino, T., Sasaki, Y., Imaoka, S., Murata, M., Shimano, T., Yamaoka, Y., and Nakamura, Y. AXIN1 mutations in hepatocellular carcinomas, and growth suppression in cancer cells by virus-mediated transfer of AXIN1, Nat Genet. *24:* 245-250, 2000.

3. Clark, E. A., Golub, T. R., Lander, E. S., and Hynes, R. O. Genomic analysis of metastasis reveals an essential role for RhoC, Nature. *406:* 532-535, 2000.

4. Okabe, H., Satoh, S., Kato, T., Kitahara, O., Yanagawa, R., Yamaoka, Y., Tsunoda, T., Furukawa, Y., and Nakamura, Y. Genome-wide analysis of gene expression in human hepatocellular carcinomas using cDNA microarray: identification of genes involved in viral carcinogenesis and tumor progression, Cancer Res. *61:* 2129-2137, 2001.

5. Jackson, T. R., Kearns, B. G., and Theibert, A. B. Cytohesins and centaurins: mediators of PI 3-kinase-regulated Arf signaling, Trends Biochem Sci. *25:* 489-495, 2000.

6. Cukierman, E., Huber, I., Rotman, M., and Cassel, D. The ARF1 GTPase-activating protein: zinc finger motif and Golgi complex localization, Science. *270:* 1999-2002, 1995.

7. Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Bloomfield, C. D., and Lander, E. S. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring, Science. *286:* 531-537, 1999.

8. Randazzo, P. A., Andrade, J., Miura, K., Brown, M. T., Long, Y. Q., Stauffer, S., Roller, P., and Cooper, J. A. The Arf GTPase-activating protein ASAP1 regulates the actin cytoskeleton, Proc Natl Acad Sci U S A. *97:* 4011-4016, 2000.

9. Cerione, R. A., and Zheng, Y. Curr. Opin. Cell. Biol. *8:* 216-222, 1996.

Chardin, P., Paris, S., Antonny, B., Robineau, S., Beraud-Dufour, S., Jackson, C. L., and Chabre, M. A human exchange factor for ARF contains Sec7- and pleckstrin-homology domains. Nature. *384:* 481-484, 1996.

10. Ren, R., Mayer, B. J., Cicchetti, P., and Baltimore, D. Identification of a ten-amino acid proline-rich SH3 binding site. Science. *259:* 1157-1161,1993.

11. Moss, J., Vaughan, M. Structure and function of ARF proteins: activators of cholera toxin and critical components of intracellular vesicular transport processes. J Biol Chem. *270:* 12327-12330, 1995.

12. Frank RC, Sun X, Berguido FJ, Jakubowiak A, Nimer SD. The t(8;21) fusion protein, AML1/ETO, transforms NIH3T3 cells and activates AP-1. Oncogene. *18*: 1701-1710, 1999 .

13. Gelmetti V, Zhang J, Fanelli M, Minucci S, Pelicci PG, Lazar MA. Aberrant recruitment of the nuclear receptor corepressor-histone deacetylase complex by the acute myeloid leukemia fusion partner ETO. Mol Cell Biol. *18*: 7185-7191, 1998.

14. Anderson SF, Schlegel BP, Nakajima T, Wolpin ES, Parvin JD. BRCA1 protein is linked to the RNA polymerase II holoenzyme complex via RNA helicase A. Nat Genet. *19*: 254-256, 1998.

15. Wolff T, Rubin GM. Strabismus, a novel gene that regulates tissue polarity and cell fate decisions in Drosophila. Development *125*:1149-1159, 1998.

16. Kibar Z, Vogan KJ, Groulx N, Justice MJ, Underhill DA, Gros P. Ltap, a mammalian homolog of Drosophila Strabismus/Van Gogh, is altered in the mouse neural tube mutant Loop-tail. Nat Genet. *28*: 251-255, 2001.

17. Mochizuki, N., Cho, G., Wen, B., Insel, P. A. Identification and cDNA cloning of a novel human mosaic protein, LGN, based on interaction with $G_{\alpha i2}$. Gene. *181:* 39-43, 1996.

18. Hermouet, S., Merendino, J. J. Jr., Gutkind, J. S., Spiegel, A. M. Activating and inactivating mutations of the alpha subunit of Gi2 protein have opposite effects on proliferation of NIH 3T3 cells. Proc Natl Acad Sci USA. *88:* 10455-10459, 1991.

20. Pace, A. M., Wong, Y. H., Bourne, H. R. A mutant alpha subunit of Gi2 induces neoplastic transformation of Rat-1 cells. Proc Natl Acad Sci USA. *88:* 7031-7035, 1991.205.

21. Lyons, J., Landis, C. A., Harsh, G., Vallar, L., Grunewald, K., Feichtinger, H., Duh, Q. Y., Clark, O. H., Kawasaki, E., Bourne, H. R., et al. Two G protein oncogenes in human endocrine tumors. Science. *249:* 655-659, 1990.

Figure Legends

Figure 1. Expression of *B9362* in HCCs a) Relative expression ratio (cancer/non-cancer) of *B9362* in primary 20 HCCs examined by cDNA microarray. Its up-regulated expression (Cy3:Cy5 intensity ratio, >2.0) was observed in 11 of the 12 HCCs that passed through the cutoff filter (both Cy3 and Cy5 signals greater than 25,000). b) Expression of *B9362* was analyzed by semi-quantitative RT-PCR using additional 11 HCC cases. Expression of *GAPDH* served as an internal control.

Figure 2. a) Northern blot analysis of *DDEFL1* in various human tissues. The transcript of *B9362* is approximately 4-kb by size. b) The structure of *DDEFL1*. The DDEFL1 cDNA consists of 4,050 nucleotides with an ORF of 2.712 nucleotides and is composed of 25 exons. c) Similarity between the expected DDEFL1 protein and members of ArfGAP family. d) Identity between the amino acid sequence of the ArfGAP motif in DDEFL1 and that in DDEF2. The arrows indicate a $CXXCX_{16}CXXC$ motif, representing a zinc finger structure essential to GAP activity.

Figure 3. Subcellular localization of DDEFL1. a) The western blot analysis showed that cMyc-tagged DDEFL1 protein was expressed in COS7 cells transfected with pcDNA-DDEFL1-myc plasmid. b) Immunocytochemistry of the cells suggested that cMyc-tagged DDEFL1 protein localized in the cytoplasm.

Figure 4. Growth-promoting effect of DDEFL1. a) Colony formation assays show that DDEFL1 promotes cell growth in NIH3T3, SNU423, and Alexander cells. b)NIH3T3-DDEFL1 cells stably expressed exogeneous DDEFL1. c) Growth of NIH3T3-DDEFL1 cells stably expressing exogeneous DDEFL1 was statistically higher than that of mock (NIH3T3-LacZ) cells in culture media containing 10% FBS ($P<0.05$). d) Growth of NIH3T3-DDEFL1 cells was statistically higher than that of mock cells in culture media containing 0.1% FBS ($P<0.01$).

Figure 5. Growth suppression by antisense S-oligonucleotides designated to suppress *DDEFL1* in SNU475 cells. a) Designation of antisense S-oligonucleotides. Reduced expression of *DDEFL1* by the transfection of AS1 or AS5 antisense S-oligonucleotides. b)AS1 and AS5 suppressed growth of SNU475 cells.

Figure 6. Expression of *A6681* in HCCs a) Relative expression ratios (cancer/non-cancer) of A6681 in primary 20 HCCs examined by cDNA microarray. Its up-regulated expression (Cy3:Cy5 intensity ratio, >2.0) was observed in 11 of the 12 HCCs that passed through the cutoff filter (both Cy3 and Cy5 signals greater than 25,000). B) Expression of *A6681* was analyzed by semi-quantitative RT-PCR using additional 10 HCC cases. T, tumor tissue; N, normal tissue. Expression of *GAPDH* served as an internal control.

Figure 7. a) Multiple-tissue northern blot analysis of *ZNFN3A1* in various human tissues. b) Predicted protein structure of ZNFN3A1.

Figure 8. Subcellular localization of ZNFN3A1 a) SNU475 cells were transfected with pEGFP- ZNFN3A1. Nuclei were counter-stained with DAPI. b) SNU475 cells transfected with pFLAG-ZNFN3A1 were stained with anti-Flag monoclonal antibody and visualized by Rhodamine conjugated secondary anti-mouse IgG antibody. Nuclei were counter-stained with DAPI.

Figure 9. Effect of *ZNFN3A1* on cell growth. a) Colony formation assay of *ZNFN3A1* in NIH3T3 cells. b) Number of colonies counted by electric densitometry. c) Expression of *ZNFN3A1* in mock, NIH3T3-antisense ZNFN3A1, and NIH3T3-ZNFN3A1 cells d) Cell growth of mock, NIH3T3-antisense ZNFN3A1, and NIH3T3-ZNFN3A1 cells.

Figure 10 Growth suppressive effect of antisense S-oligonucleotides designated to suppress *ZNFN3A1*. a) Expression of *ZNFN3A1* in SNU475 cells treated with either control (SE) or antisense (AS) oligonucleotides for 12 h. b) S-

oligonucleotides (AS) suppressed growth of SNU423 cells. c) Analysis of cell viability by MTT assay. d) FACS analysis of Huh7 cells treated with control (SE) or antisense (AS) oligonucleotides.

Figure 11. Expression of *VANGL1* in HCCs a) Relative expression ratios (cancer/non-cancer) of *VANGL1* in primary 20 HCCs examined by cDNA microarray. Its up-regulated expression (Cy3:Cy5 intensity ratio, >2.0) was observed in 10 of the 12 HCCs that passed through the cutoff filter (both Cy3 and Cy5 signals greater than 25,000). b) Expression of *A6681* was analyzed by semi-quantitative RT-PCR using additional 10 HCC cases. T, tumor tissue; N, normal tissue. Expression of *GAPDH* served as an internal control.

Figure 12. a) Multiple-tissue northern blot analysis of *VANGL1* in various human tissues. b) Predicted protein structure of VANGL1.

Figure 13. Subcellular localization of VANGL1 a) SNU475 cells transfected with pcDNA3.1-myc/His-LGN were stained with mouse anti-myc monoclonal antibody and visualized by Rhodamine conjugated secondary anti-mouse IgG antibody. Nuclei were counter-stained with DAPI.

Figure 14 Growth suppressive effect of antisense S-oligonucleotides designated to suppress *VANGL1*. a) Expression of *VANGL1* in SNU475 cells treated with either control or antisense oligonucleotides for 12 h. b) S-oligonucleotides suppressed growth of SNU423 cells. c) Analysis of cell viability by MTT assay. d) FACS analysis of cells treated with sense or antisense oligonucleotides.

Figure 15 LGN gene expression of HCCs compared with their corresponding non-cancerous liver tissues. a) Relative expression ratios (cancer/non-cancer) of *LGN* in primary 20 HCCs studied by cDNA microarray. Its up-regulated expression (Cy3:Cy5 intensity ratio, >2.0) was observed in 10 of the 12 HCCs that passed through the cutoff filter (both Cy3 and Cy5 signals greater than 25,000). b) Expression of *LGN* was analyzed by semi-quantitative RT-PCR using additional ten HCCs. RT-PCR experiments were performed with a *LGN*-specific primer set using

RNA extracted from HCCs. Expression of *GAPDH* served as an internal control. T, tumor tissue; N, normal tissue.

Figure 16. Genomic structure of *LGN*.

Figure 17. Subcellular localization of LGN. COS7 cells transfected with pcDNA3.1-myc/His-LGN were stained with mouse anti c-myc antibody and visualized by Rhodamine conjugated secondary anti-mouse IgG antibody. Nuclei were counter-stained with DAPI.

Figure 18. Growth-promoting effect of *LGN*. a) Colony formation assays show that *LGN* promotes cell growth in NIH3T3, SNU423, Alexander, and SNU423 cells. b) Growth of NIH3T3-LGN cells stably expressing exogeneous *LGN* was higher than that of mock (NIH3T3-LacZ) cells in culture media containing 10% FBS.

Figure 19. Growth suppression by antisense S-oligonucleotides designated to suppress *LGN* expression in human hepatoma SNU423 cells. a) Reduced expression of *LGN* by the transfection of antisense S-oligonucleotides #4. b) S-oligonucleotides #4 suppressed growth of SNU423 cells.

Sequence_Table

SEQ ID 1
DDEFL1(nucleotide)
GTGCCCCCCGCGCTCCGCTCCGGCAGCTCCACGCTCGCGCCCGCCATGCCGGAGCA
GTTCAGCGTCGCCGAGTTCCTGGCCGTCACCGCGGAGGACCTCAGCTCCCCGGCTG
GGGCCGCCGCCTTCGCCGCCAAGATGCCCCGGTACCGAGGGGCGGCGCTGGCGCG
GGAGGAGATCTTGGAAGGAGACCAAGCCATCCTGCAGAGAATAAAGAAGGCTGTGCG
GGCAATCCATAGCTCCGGCCTTGGCCATGTGGAGAATGAAGAGCAGTACCGAGAGGC
CGTGGAATCCTTAGGCAACAGCCACCTGTCCCAGAACAGCCATGAGCTGTCCACAGG
CTTCCTAAACTTGGCCGTGTTCACCCGCGAGGTTGCTGCGCTCTTCAAGAACCTGATT
CAGAACTTGAACAACATTGTCTCTTTCCCCCTGGACAGTCTGATGAAGGGGCAGCTGA
GGGACGGTCGACAGGATTCCAAAAAACAGCTGGAGAAGGCATGGAAGGACTATGAAG
CCAAAATGGCCAAGCTGGAGAAGGAGCGCGATCGGGCCAGGGTGACAGGAGGGATC
CCTGGGGAGGTGGCCCAGGACATGCAGAGAGAGCGGCGCATCTTCCAGCTGCACAT
GTGTGAGTATCTGCTCAAAGCCGGGGAGAGCCAGATGAAGCAAGGTCCTGACTTCCTT
CAGAGCCTCATCAAGTTCTTCCACGCCCAGCACAACTTTTTCCAAGATGGCTGGAAGG
CTGCCCAGAGCCTGTTCCCCTTCATCGAGAAGCTGGCGGCCTCAGTACATGCACTCCA
TCAGGCCCAGGAGGACGAGCTACAGAAGCTGACCCAGCTCCGGGACTCCCTCCGAG
GGACACTGCAGCTTGAGAGCAGAGAGGAACACCTGAGCCGGAAGAACTCAGGATGTG
GCTATAGCATCCACCAGCACCAAGGCAACAAGCAGTTTGGGACGGAGAAAGTGGGCT
TTCTATACAAGAAAAGTGACGGAATTCGAAGAGTCTGGCAGAAAAGGAAGTGTGGAGT
CAAGTATGGCTGCCTGACCATCTCACACAGCACGATAAACCGGCCCCCGGTGAAGCTG
ACCCTGCTGACGTGCCAAGTGAGGCCAAACCCTGAGGAGAAAAAGTGCTTCGACCTG
GTGACCCACAACCGGACGTACCACTTTCAGGCAGAGGACGAGCACGAGTGTGAGGC
GTGGGTGTCAGTGTTGCAGAACAGCAAGGACGAAGCCCTGAGCAGCGCCTTCCTCG
GGGAGCCCAGCGCTGGCCCGGGGTCCTGGGGGTCCGCCGGCCATGATGGGGAGCC
GCACGACCTCACAAAGCTGCTCATCGCGGAGGTGAAGAGCAGGCCTGGGAATAGCCA
GTGCTGCGACTGCGGGGCTGCAGACCCCACGTGGCTCAGCACCAACCTGGGCGTGC
TCACCTGCATCCAGTGCTCGGGCGTCCACCGCGAACTGGGCGTGCGCTTTTCGCGCA
TGCAGTCACTCACCTTGGACCTGCTGGGCCCCTCCGAGTTGTTGCTGGCCTTGAACAT
GGGAAACACGAGCTTCAATGAGGTCATGGAGGCCCAGCTACCCTCACACGGCGGCCC
TAAACCCTCAGCTGAGAGTGACATGGGCACCCGCAGGGACTACATTATGGCCAAGTAT
GTGGAGCATAGGTTTGCACGCCGGTGCACACCTGAGCCTCAGCGACTCTGGACAGCC
ATTTGCAACAGGGACCTCCTGTCGGTACTGGAGGCCTTTGCCAATGGGCAGGACTTTG
GACAGCCGCTGCCAGGGCCTGATGCACAGGCACCTGAAGAACTCGTCTTGCATTTGG
CTGTCAAAGTCGCCAACCAGGCTTCCCTGCCTCTGGTGGATTTCATCATCCAGAACGG
TGGTCACCTGGATGCCAAGGCTGCTGACGGGAACACGGCTCTGCACTACGCAGCACT
CTACAACCAGCCCGACTGCCTCAAGCTGCTGCTGAAGGGGAGAGCTTTGGTTGGCAC
AGTAAATGAAGCAGGCGAGACAGCTCTGGACATAGCCAGGAAGAAGCACCACAAGGA
GTGTGAGGAGCTGCTGGAGCAGGCCCAGGCGGGGACCTTTGCCTTCCCTCTACATGT
GGACTACTCCTGGGTAATTTCCACAGAGCCTGGCTCTGACAGTGAGGAGGATGAGGAA
GAGAAGCGCTGCTTGCTGAAGCTCCCGGCCCAGGCTCACTGGGCCAGTGGGAGGCT
GGACATCAGCAACAAGACCTATGAGACTGTCGCCAGCCTGGGAGCAGCCACCCCTCA
GGGCGAGAGTGAGGACTGTCCCCCGCCCTTGCCAGTCAAAAACTCTTCTCGGACTTT
GGTCCAAGGGTGTGCAAGACATGCCAGTGGAGATCGTTCTGAAGTCTCCAGCCTGAG
TTCAGAGGCCCCTGAGACCCCTGAGAGCCTGGGCAGTCCAGCCTCCTCCTCCAGTCT
GATGAGCCCCTTGGAACCTGGGGATCCCAGCCAAGCCCCACCCAACTCTGAAGAGGG
CCTCCGAGAGCCCCCAGGCACCTCCAGACCCAGCCTGACATCCGGGACCACCCCTTC
GGAGATGTACCTCCCCGTCAGATTCAGCTCCGAGAGCACTCGCTCCTATCGGCGGGG
GGCGCGGAGCCCTGAAGATGGTCCCTCAGCCAGGCAGCCTCTGCCCAGAAGGAACG
TGCCGGTTGGCATCACTGAAGGAGATGGCTCAAGGACTGGGAGTCTCCCAGCAAGTT
CTGTGCAACTTTTGCAAGACTAGCTCCTTGCTGGCCCCCACATGCCCCATGCTAGGCC
CCAATGTTCAGAGCTGGGACTTGAGCTCACAAAACTGGGGAGCTGAGACATTTGTTCT
CTTGGATCTCACTCTCTCTGTCCCTTGTGCCTCTGTAGCTGGCCTTCTTCCTGCCACAG
GCCATGCCTCTACCAAGGACACATGGCCTTTCCCTGTTAGGGCTGATGGCGGTTCTTT

CCTATCTCATTACCCGCTAGGGGCCTGGGAGCCCTGTGGCTGGATCTGAGTGCTCCTG
AGCTGGCTTCAGCTGCAGAACTCTCAGTCCCTCATCAGATCGAGACTCTATTTCCCCC
GTCAGTCTGGGGGCTTCACAAGGGCAGGAGAGCCCTCCATCACTGACTTCCAGATCA
GGGACCCTGCCAAGTAGGGACTGTCTTCTCAGCCAGCCATTTATTAGTCTAATATTCCT
TCACTAAATTCCAACTCTATGTCTGGACCTGTGTTAGGCACTTCAGATACCACACGAGTA
AGACAAGGGCCCTGCAGGGGTGGTCCTTTGGTGGAAAGCTGGTCTTAAGGGTTGGGC
TTGGGAATAGGCAGGGTCAGATTCCAGGGCATGGCTCTGGACTCAGCTGGTTTATACC
TATATGACCATTACAGTTGTCTACAGATCACATCCATTCTGGCTGGTCAACATGCATGCT
GTACTGGCTGTTAAATAAAAATATTCTGAATGTCACTCCTTTTGAGGGACAGCACAGCCT
TCCCTAGGCATTCTCCTATATTCCCCAGCCAAATTGTAGAGTCAGATGCACCCACATTTG
CCTGTGTCCTTGATTTAGCAGGAAGGAAAGGAATAGTCGGGGTTGATGGATGCCCACT
TCTCTTCTCTTTCTCTTGGTCAACTCAGGAGCCTTTTAGTCTGAGGGAATGGAGAGGCA
AAGAAAGAAGGGAGAGTAATAGAATTGGGAGGGCAGAGACTTAAGGGTTCTGCTTCCC
AGCCCTAGAAATTCTATCATTGCTCAGCCCCAATGAGAAAGCAGATACACCTAAGCCAT
CATCAACCACTAACATCTCAACTTGCCAGTTGCTGGGTGCTGGGCCCTGGCAGGAATG
GGCCAAGCCAAGCAGGGGAGACTAGAGAGCACCAATGGCCAACACAGCTGCCTGGC
TGGGGAGGCTGTGCTGTTTCCCCTGGAGACCTGACTGGTCTGTGGTTCCCACAGGAA
CAGGGTTGTCTTTTGAGCCCCCAGTGTCTGGTTTCATTCATCTCAGACTTGTTATTTCA
CTCATCTCTAATAAAGGATTGGGGGGTCAGTT

SEQ ID 2
DDEFL1(amino acid)
MPEQFSVAEFLAVTAEDLSSPAGAAAFAAKMPRYRGAALAREEILEGDQAILQRIKKAVRAI
HSSGLGHVENEEQYREAVESLGNSHLSQNSHELSTGFLNLAVFTREVAALFKNLIQNLNNI
VSFPLDSLMKGQLRDGRQDSKKQLEKAWKDYEAKMAKLEKERDRARVTGGIPGEVAQD
MQRERRIFQLHMCEYLLKAGESQMKQGPDFLQSLIKFFHAQHNFFQDGWKAAQSLFPFIE
KLAASVHALHQAQEDELQKLTQLRDSLRGTLQLESREEHLSRKNSGCGYSIHQHQGNKQF
GTEKVGFLYKKSDGIRRVWQKRKCGVKYGCLTISHSTINRPPVKLTLLTCQVRPNPEEKKC
FDLVTHNRTYHFQAEDEHECEAWVSVLQNSKDEALSSAFLGEPSAGPGSWGSAGHDGE
PHDLTKLLIAEVKSRPGNSQCCDCGAADPTWLSTNLGVLTCIQCSGVHRELGVRFSRMQS
LTLDLLGPSELLLALNMGNTSFNEVMEAQLPSHGGPKPSAESDMGTRRDYIMAKYVEHRF
ARRCTPEPQRLWTAICNRDLLSVLEAFANGQDFGQPLPGPDAQAPEELVLHLAVKVANQA
SLPLVDFIIQNGGHLDAKAADGNTALHYAALYNQPDCLKLLLKGRALVGTVNEAGETALDIA
RKKHHKECEELLEQAQAGTFAFPLHVDYSWVISTEPGSDSEEDEEEKRCLLKLPAQAHWA
SGRLDISNKTYETVASLGAATPQGESEDCPPPLPVKNSSRTLVQGCARHASGDRSEVSSL
SSEAPETPESLGSPASSSSLMSPLEPGDPSQAPPNSEEGLREPPGTSRPSLTSGTTPSEM
YLPVRFSSESTRSYRRGARSPEDGPSARQPLPRRNVPVGITEGDGSRTGSLPASSVQLLQ
D

SEQ ID 3
ZNFN3A1(nucleotide)
GTGCGCGCAGGGCGCAGGCGCGCGGGTCCCGGCAGCCCGTGAGACGCCCGCTGCT
GGACGCGGGTAGCCGTCTGAGGTGCCGGAGCTGCGGGAGGATGGAGCCGCTGAAG
GTGGAAAAGTTCGCAACCGCCAACAGGGGAAACGGGCTGCGCGCCGTGACCCCGCT
GCGCCCCGGAGAGCTACTCTTCCGCTCGGATCCCTTGGCGTACACGGTGTGCAAGGG
GAGTCGTGGCGTCGTCTGCGACCGCTGCCTTCTCGGGAAGGAAAAGCTGATGCGATG
CTCTCAGTGCCGCGTCGCCAAATACTGTAGTGCTAAGTGTCAGAAAAAAGCTTGGCCA
GACCACAAGCGGGAATGCAAATGCCTTAAAAGCTGCAAACCCAGATATCCTCCAGACT
CCGTTCGACTTCTTGGCAGAGTTGTCTTCAAACTTATGGATGGAGCACCTTCAGAATCA
GAGAAGCTTTACTCATTTTATGATCTGGAGTCAAATATTAACAAACTGACTGAAGATAAG
AAAGAGGGCCTCAGGCAACTCGTAATGACATTTCAACATTTCATGAGAGAAGAAATACA
GGATGCCTCTCAGCTGCCACCTGCCTTTGACCTTTTTGAAGCCTTTGCAAAAGTGATCT
GCAACTCTTTCACCATCTGTAATGCGGAGATGCAGGAAGTTGGTGTTGGCCTATATCCC
AGTATCTCTTTGCTCAATCACAGCTGTGACCCCAACTGTTCGATTGTGTTCAATGGGCC
CCACCTCTTACTGCGAGCAGTCCGAGACATCGAGGTGGGAGAGGAGCTCACCATCTG
CTACCTGGATATGCTGATGACCAGTGAGGAGCGCCGGAAGCAGCTGAGGGACCAGTA

CTGCTTTGAATGTGACTGTTTCCGTTGCCAAACCCAGGACAAGGATGCTGATATGCTAA
CTGGTGATGAGCAAGTATGGAAGGAAGTTCAAGAATCCCTGAAAAAAATTGAAGAACTG
AAGGCACACTGGAAGTGGGAGCAGGTTCTGGCCATGTGCCAGGCGATCATAAGCAGC
AATTCTGAACGGCTTCCCGATATCAACATCTACCAGCTGAAGGTGCTCGACTGCGCCAT
GGATGCCTGCATCAACCTCGGCCTGTTGGAGGAAGCCTTGTTCTATGGTACTCGGACC
ATGGAGCCATACAGGATTTTTTTCCCAGGAAGCCATCCCGTCAGAGGGGTTCAAGTGA
TGAAAGTTGGCAAACTGCAGCTACATCAAGGCATGTTTCCCCAAGCAATGAAGAATCTG
AGACTGGCTTTTGATATTATGAGAGTGACACATGGCAGAGAACACAGCCTGATTGAAGA
TTTGATTCTACTTTTAGAAGAATGCGACGCCAACATCAGAGCATCCTAAGGGAACGCAG
TCAGAGGGAAATACGGCGTGTGTCTTTGTTGAATGCCTTATTGAGGTCACACACTCTAT
GCTTTGTTAGCTGTGTGAACCTCTCTTATTGGAAATTCTGTTCCGTGTTTGTGTAGGTAA
ATAAAGGCAGACATGGTTTGCAAACCACAAGAATCATTAGTTGTAGAGAAGCACGATTAT
AATAAATTCAAAACATTTGGTTGAGGATGCCAAAAAAAAAAAAAAAAAAAA

SEQ ID 4
ZNFN3A1(amino acid)
MEPLKVEKFATANRGNGLRAVTPLRPGELLFRSDPLAYTVCKGSRGVVCDRCLLGKEKLM
RCSQCRVAKYCSAKCQKKAWPDHKRECKCLKSCKPRYPPDSVRLLGRVVFKLMDGAPS
ESEKLYSFYDLESNINKLTEDKKEGLRQLVMTFQHFMREEIQDASQLPPAFDLFEAFAKVIC
NSFTICNAEMQEVGVGLYPSISLLNHSCDPNCSIVFNGPHLLLRAVRDIEVGEELTICYLDML
MTSEERRKQLRDQYCFECDCFRCQTQDKDADMLTGDEQVWKEVQESLKKIEELKAHWK
WEQVLAMCQAIISSNSERLPDINIYQLKVLDCAMDACINLGLLEEALFYGTRTMEPYRIFFP
GSHPVRGVQVMKVGKLQLHQGMFPQAMKNLRLAFDIMRVTHGREHSLIEDLILLLEECDA
NIRAS

SEQ ID 5
VANGL1
CTCGCTCACAAAAAATTGAGCCGGCCCTGGAGGCCTGGGGGGCGAGTCCGGTTGCG
CCTCGGAGAGCGCAACAGGCAGAATTTGTTCCTGTTGAAGAGTGGCTCCTCTTCTAAT
TTCCAGACTCCTTGAGGTTTTAGGAGTCTGGTAGGTGAAATTTTCTACCTCTAAGGAGA
AACAGTACCTGCTCCTTCCTCAAGCGCAAGCCCTCCATTGCTATGGATACCGAATCCAC
TTATTCTGGATATTCTTACTATTCAAGTCATTCGAAAAAATCTCACAGACAAGGGGAAAG
AACTAGAGAGAGACACAAGTCACCCCGGAATAAAGACGGCAGAGGGTCAGAAAAGTC
TGTCACCATTCAACCTCCCACTGGAGAGCCCCTGTTGGGAAATGATTCTACTCGGACA
GAGGAAGTTCAGGATGACAACTGGGGAGAGACCACCACGGCCATCACAGGCACCTCG
GAGCACAGCATATCCCAAGAGGACATTGCCAGGATCAGCAAGGACATGGAGGACAGC
GTGGGGCTGGATTGCAAACGCTACCTGGGCCTCACCGTCGCCTCTTTTCTTGGACTTC
TAGTTTTCCTCACCCCTATTGCCTTCATCCTTTTACCTCCGATCCTGTGGAGGGATGAG
CTGGAGCCTTGTGGCACAATTTGTGAGGGGCTCTTTATCTCCATGGCATTCAAACTCCT
CATTCTGCTCATAGGGACCTGGGCACTTTTTTTCCGCAAGCGGAGAGCTGACATGCCA
CGGGTGTTTGTGTTTCGTGCCCTTTTGTTGGTCCTCATCTTTCTCTTTGTGGTTTCCTAT
TGGCTTTTTTACGGGGTCCGCATTTTGGACTCTCGGGACCGGAATTACCAGGGCATTG
TGCAATATGCAGTCTCCCTTGTGGATGCCCTCCTCTTCATCCATTACCTGGCCATCGTC
CTGCTGGAGCTCAGGCAGCTGCAGCCCATGTTCACGCTGCAGGTGGTCCGCTCCACC
GATGGCGAGTCCCGCTTCTACAGCCTGGGACACCTGAGTATCCAGCGAGCAGCATTG
GTGGTCCTAGAAAATTACTACAAAGATTTCACCATCTATAACCCAAACCTCCTAACAGCC
TCCAAATTCCGAGCAGCCAAGCATATGGCCGGGCTGAAAGTCTACAATGTAGATGGCC
CCAGTAACAATGCCACTGGCCAGTCCCGGGCCATGATTGCTGCAGCTGCTCGGCGCA
GGGACTCAAGCCACAACGAGTTGTATTATGAAGAGGCCGAACATGAACGGCGAGTAAA
GAAGCGGAAAGCAAGGCTGGTGGTTGCAGTGGAAGAGGCCTTCATCCACATTCAGCG
TCTCCAGGCTGAGGAGCAGCAGAAAGCCCCAGGGGAGGTGATGGACCCTAGGGAGG
CCGCCCAGGCCATTTTCCCCTCCATGGCCAGGGCTCTCCAGAAGTACCTGCGCATCAC
CCGGCAGCAGAACTACCACAGCATGGAGAGCATCCTGCAGCACCTGGCCTTCTGCAT
CACCAACGGCATGACCCCCAAGGCCTTCCTAGAACGGTACCTCAGTGCGGGCCCCAC
CCTGCAATATGACAAGGACCGCTGGCTCTCTACACAGTGGAGGCTTGTCAGTGATGAG
GCTGTGACTAATGGATTACGGGATGGAATTGTGTTCGTCCTTAAGTGCTTGGACTTCAG

CCTCGTAGTCAATGTGAAGAAAATTCCATTCATCATACTCTCTGAAGAGTTCATAGACCC
CAAATCTCACAAATTTGTCCTTCGCTTACAGTCTGAGACATCCGTTTAAAAGTTCTATATT
TGTGGCTTTATTAAAAAAAAAAGAAAAATATATAGAGAGATATGCAAAAAAAATAAAAGAC
AAAAACAAAAAAAAAA

SEQ ID 6
VANGL1(amino acid)
MDTESTYSGYSYYSSHSKKSHRQGERTRERHKSPRNKDGRGSEKSVTIQPPTGEPLLGN
DSTRTEEVQDDNWGETTTAITGTSEHSISQEDIARISKDMEDSVGLDCKRYLGLTVASFLGL
LVFLTPIAFILLPPILWRDELEPCGTICEGLFISMAFKLLILLIGTWALFFRKRRADMPRVFVFR
ALLLVLIFLFVVSYWLFYGVRILDSRDRNYQGIVQYAVSLVDALLFIHYLAIVLLELRQLQPMF
TLQVVRSTDGESRFYSLGHLSIQRAALVVLENYYKDFTIYNPNLLTASKFRAAKHMAGLKVY
NVDGPSNNATGQSRAMIAAAARRRDSSHNELYYEEAEHERRVKKRKARLVVAVEEAFIHIQ
RLQAEEQQKAPGEVMDPREAAQAIFPSMARALQKYLRITRQQNYHSMESILQHLAFCITN
GMTPKAFLERYLSAGPTLQYDKDRWLSTQWRLVSDEAVTNGLRDGIVFVLKCLDFSLVVN
VKKIPFIILSEEFIDPKSHKFVLRLQSETSV

SEQ ID 7
LGN(nucleotide)
GGCACGAGGAAGAATCAGGAGCTTAGGATGTATTAACACCAACTCATTAATATACTAACC
GGACAATGTTCTACAAACAATTCTACATTGTAAAGGACTGGATTGGCACAAAATAAAATA
ATTTTATTTTATTCAGCTTATAATATGACTCGATGGAGGAAAATTTGATAAGCATGAGAGA
AGACCATTCTTTTCATGTTCGTTACAGAATGGAAGCTTCTTGCCTAGAGCTGGCCTTGG
AAGGGGAACGTCTATGTAAATCAGGAGACTGCCGCGCTGGCGTGTCATTCTTTGAAGC
TGCAGTTCAAGTTGGAACTGAAGACCTAAAAACACTTAGCGCTATTTACAGCCAGTTGG
GCAATGCTTATTTCTATTTGCATGATTATGCCAAAGCATTAGAATATCACCATCATGATTTA
ACCCTTGCAAGGACTATTGGAGACCAGCTGGGGGAAGCGAAAGCTAGTGGTAATCTG
GGAAACACCTTAAAAGTTCTTGGGAATTTTGACGAAGCCATAGTTTGTTGTCAGCGACA
CCTAGATATTTCCAGAGAGCTTAATGACAAGGTGGGAGAAGCAAGAGCACTTTACAATC
TTGGGAATGTGTATCATGCCAAAGGGAAAAGTTTTGGTTGCCCTGGTCCCCAGGATGT
AGGAGAATTTCCAGAAGAAGTGAGAGATGCTCTGCAGGCAGCCGTGGATTTTTATGAG
GAAAACCTATCATTAGTGACTGCTTTGGGTGACCGAGCGGCACAAGGACGTGCCTTTG
GAAATCTTGGAAACACACATTACCTCCTTGGCAACTTCAGGGATGCAGTTATAGCTCAT
GAGCAGCGTCTCCTTATTGCAAAAGAATTTGGAGATAAAGCAGCTGAAAGAAGAGCATA
TAGCAACCTTGGAAATGCATATATATTTCTTGGTGAATTTGAAACTGCCTCGGAATACTAC
AAGAAGACACTACTGTTGGCCCGACAGCTTAAAGACCGAGCTGTAGAAGCACAGTCTT
GTTACAGTCTTGGAAATACATATACTTTACTTCAAGACTATGAAAAGGCCATTGATTATCA
TCTGAAGCACTTAGCAATTGCTCAAGAGCTGAATGATAGAATTGGTGAAGGAAGAGCAT
GTTGGAGCTTAGGAAATGCATACACAGCACTAGGAAATCATGATCAAGCAATGCATTTT
GCTGAAAAGCACTTGGAAATTTCAAGAGAGGTTGGGGATAAAAGTGGTGAACTAACAG
CACGACTTAATCTCTCAGACCTTCAAATGGTTCTTGGTCTGAGCTACAGCACAAATAAC
TCCATAATGTCTGAAAATACTGAAATTGATAGCAGTTTGAATGGTGTACTCCCCAAGTTG
GGACGCCGGCATAGTATGGAAAATATGGAACTTATGAAGTTAACACCAGAAAAGGTACA
GAACTGGAACAGTGAAATTCTTGCTAAGCAAAAACCTCTTATTGCCAAACCTTCTGCAA
AGCTACTCTTTGTCAACAGACTGAAGGGGAAAAAATACAAAACGAATTCCTCCACTAAA
GTTCTCCAAGATGCCAGTAATTCTATTGACCACCGAATTCCAAATTCTCAGAGGAAAATC
AGTGCAGATACTATTGGAGATGAAGGGTTCTTTGACTTATTAAGCCGATTTCAAAGCAAT
AGGATGGATGATCAGAGATGTTGCTTACAAGAAAAGAACTGCCATACAGCTTCAACAAC
AACTTCTTCCACTCCCCCTAAAATGATGCTAAAAACATCATCTGTTCCTGTGGTATCCCC
CAACACGGATGAGTTTTTAGATCTTCTTGCCAGCTCACAGAGTCGCCGTCTGGATGAC
CAGAGGGCTAGTTTCAGTAATTTGCCAGGGCTTCGTCTAACACAAAACAGCCAGTCGG
TACTTAGCCACCTGATGACTAATGACAACAAAGAGGCTGATGAAGATTTCTTTGACATCC
TTGTAAAATGTCAAGGATCCAGATTAGATGATCAAAGATGTGCTCCACCACCTGCTACC
ACAAAGGGTCCGACAGTACCAGATGAAGACTTTTTCAGCCTTATTTTACGGTCCCAGGG
AAAGAGAATGGATGAACAGAGAGTTCTTTTACAAAGAGATCAAAACAGAGACACTGACT
TTGGGCTAAAGGACTTTTTGCAAAATAATGCTTTGTTGGAGTTTAAAAATTCAGGGAAAA

AATCGGCAGACCATTAGTTACTATGGATTTATTTTTTTTCCTTTCAAACACGGTAAGGAAA
CAATCTATTACTTTTTTCCTTAAAAGGAGAATTTATAGCACTGTAATACAGCTTAAAATATT
TTTAGAATGATGTAAATAGTTAA

SEQ ID 8
LGN(amino acid)
MREDHSFHVRYRMEASCLELALEGERLCKSGDCRAGVSFFEAAVQVGTEDLKTLSAIYSQ
LGNAYFYLHDYAKALEYHHHDLTLARTIGDQLGEAKASGNLGNTLKVLGNFDEAIVCCQRH
LDISRELNDKVGEARALYNLGNVYHAKGKSFGCPGPQDVGEFPEEVRDALQAAVDFYEEN
LSLVTALGDRAAQGRAFGNLGNTHYLLGNFRDAVIAHEQRLLIAKEFGDKAAERRAYSNLG
NAYIFLGEFETASEYYKKTLLLARQLKDRAVEAQSCYSLGNTYTLLQDYEKAIDYHLKHLAIA
QELNDRIGEGRACWSLGNAYTALGNHDQAMHFAEKHLEISREVGDKSGELTARLNLSDLQ
MVLGLSYSTNNSIMSENTEIDSSLNGVLPKLGRRHSMENMELMKLTPEKVQNWNSEILAK
QKPLIAKPSAKLLFVNRLKGKKYKTNSSTKVLQDASNSIDHRIPNSQRKISADTIGDEGFFDL
LSRFQSNRMDDQRCCLQEKNCHTASTTTSSTPPKMMLKTSSVPVVSPNTDEFLDLLASSQ
SRRLDDQRASFSNLPGLRLTQNSQSVLSHLMTNDNKEADEDFFDILVKCQGSRLDDQRCA
PPPATTKGPTVPDEDFFSLILRSQGKRMDEQRVLLQRDQNRDTDFGLKDFLQNNALLEFK
NSGKKSADH

CLAIMS

1. A marker for diagnosing cancer, wherein said marker is an expression product of any one of(i) the DDEFL1 gene set forth in SEQ ID NO: 1, (ii) the ZNFN3A1 gene set forth in SEQ ID NO: 3, (iii) the VANGL1 gene set forth in SEQ ID NO: 5, or (iv) the LGN gene set forth in SEQ ID NO: 7.
2. The marker according to claim 1, wherein a transcription product of the DDEFL1 gene is the protein set forth in SEQ ID NO: 2.
3. The marker according to claim 1, wherein a transcription product of the ZNFN3A1 gene is the protein set forth in SEQ ID NO: 4.
4. The marker according to claim 1, wherein a transcription product of the VANGL1 gene is the protein set forth in SEQ ID NO: 6.
5. The marker according to claim 1, wherein a transcription product of the LGN gene is the protein set forth in SEQ ID NO: 8.
6. A tool for detecting the marker according to claim 1, wherein said marker comprises (a) a nucleotide (i) complementary to the nucleotide sequence set forth in SEQ ID NO: 1, or to the complementary strand thereof, and, (ii) comprising at least 15 nucleotides, or (b) an antibody against the protein set forth in SEQ ID NO: 2.
7. A tool for detecting the marker according to claim 1, wherein said marker comprises (a) a nucleotide (i) complementary to the nucleotide sequence set forth in SEQ ID NO: 3, or to the complementary strand thereof, and, (ii) comprising at least 15 nucleotides, or (b) an

antibody against the protein set forth in SEQ ID NO: 4.

8. A tool for detecting the marker according to claim 1, wherein said marker comprises (a) a nucleotide (i) complementary to the nucleotide sequence set forth in SEQ ID NO: 5, or to the complementary strand thereof, and, (ii) comprising at least 15 nucleotides, or (b) an antibody against the protein set forth in SEQ ID NO: 6.

9. A tool for detecting the marker according to claim 1, wherein said marker comprises (a) a nucleotide (i) complementary to the nucleotide sequence set forth in SEQ ID NO: 7, or to the complementary strand thereof, and, (ii) comprising at least 15 nucleotides, or (b) an antibody against the protein set forth in SEQ ID NO: 8.

10. A method for diagnosing cancer, comprising the step of measuring the expression level of any one of (i) the DDEFL1 gene set forth in SEQ ID NO: 1, (ii) the ZNFN3A1 gene set forth in SEQ ID NO: 3, (iii) the VANGL1 gene set forth in SEQ ID NO: 5, or (iv) the LGN gene set forth in SEQ ID NO: 7.

11. A use of the DDEFL1 gene set forth in SEQ ID NO: 1, or the DDEFL1 protein set forth in SEQ ID NO: 2, as a target for developing drugs for treating cancer.

12. A use of the ZNFN3A1 gene set forth in SEQ ID NO: 3, or the ZNFN3A1 protein set forth in SEQ ID NO: 4, as a target for developing drugs for treating cancer.

13. A use of the VANGL1 gene set forth in SEQ ID NO: 5, or the VANGL1 protein set forth in SEQ ID NO: 6, as a target for developing drugs for treating cancer.

14. A use of the LGN gene set forth in SEQ ID NO: 7, or the LGN protein set forth in SEQ ID NO: 8, as a target for developing drugs for treating cancer.

15. A DNA comprising the nucleotide sequence set forth in

SEQ ID NO: 1.

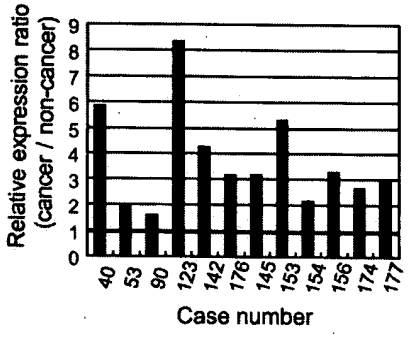
16. A protein comprising the amino acid sequence set forth in SEQ ID NO: 2.

17. A DNA comprising the nucleotide sequence set forth in SEQ ID NO: 3.

18. A protein comprising the amino acid sequence set forth in SEQ ID NO: 4.

Figure 1

a)


Relative expression ratio
(cancer / non-cancer)
0
1
2
3
4
5
6
7
8
9
40
53
90
123
142
176
145
153
154
156
174
177
Case number

b)

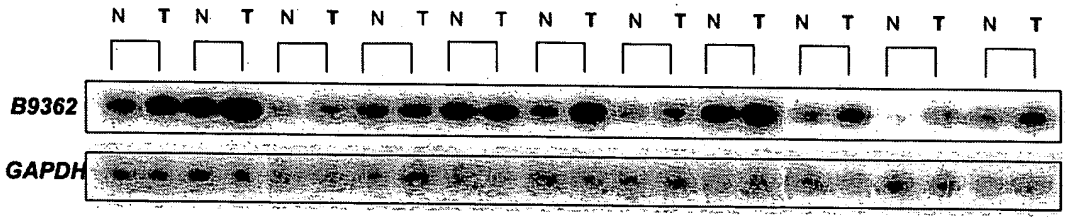

N T N T N T N T N T N T N T N T N T N T N T
B9362
GAPDH

Figure 2

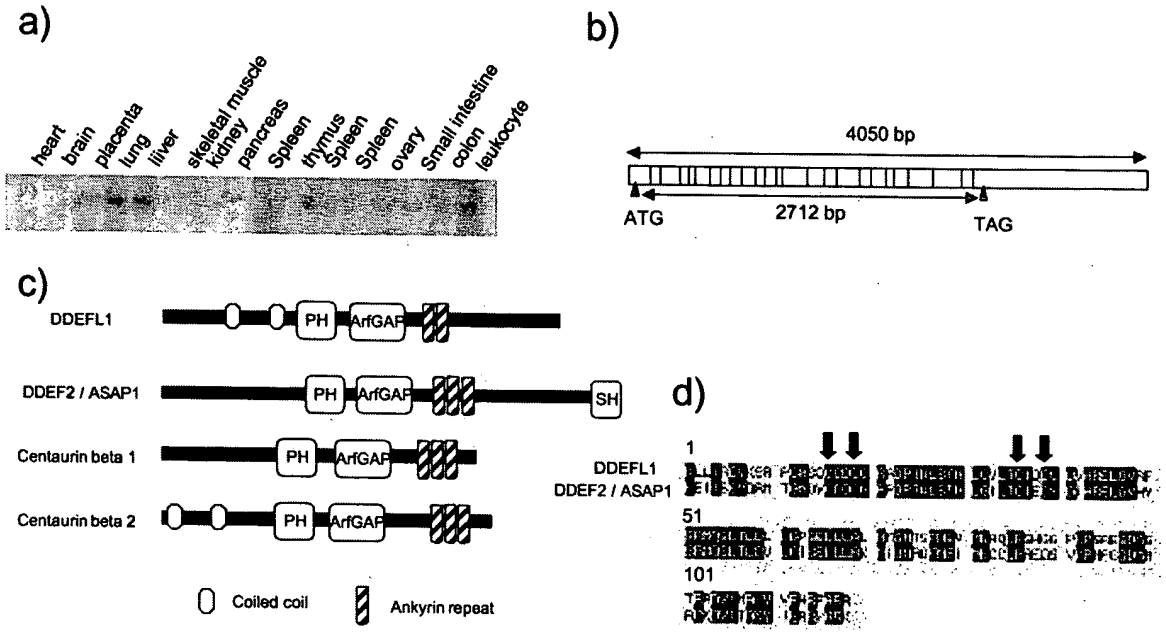

a)
heart
brain
placenta
lung
liiver
skeletal muscle
kidney
pancreas
Spleen
thymus
Spleen
Spleen
ovary
Small intestine
colon
leukocyte
b)
4050 bp
ATG
2712 bp
TAG
c)
DDEFL1
PH
ArfGAP
DDEF2 / ASAP1
SH
Centaurin beta 1
Centaurin beta 2
Coiled coil
Ankyrin repeat
d)
1
DDEFL1
DDEF2 / ASAP1
51
101

Figure 3

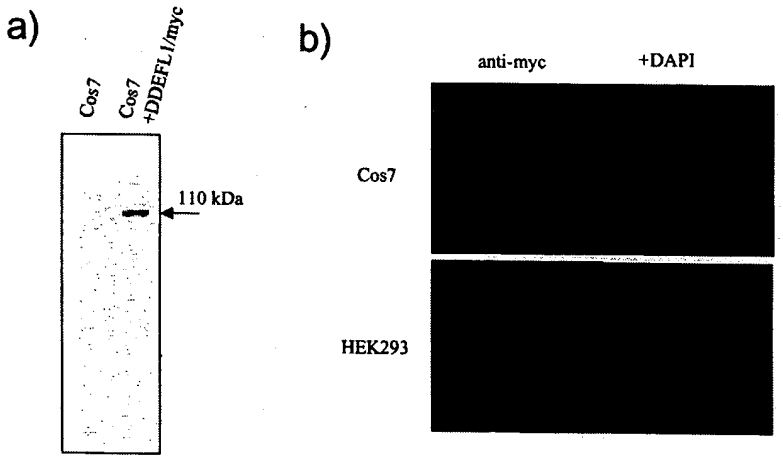


Figure 4

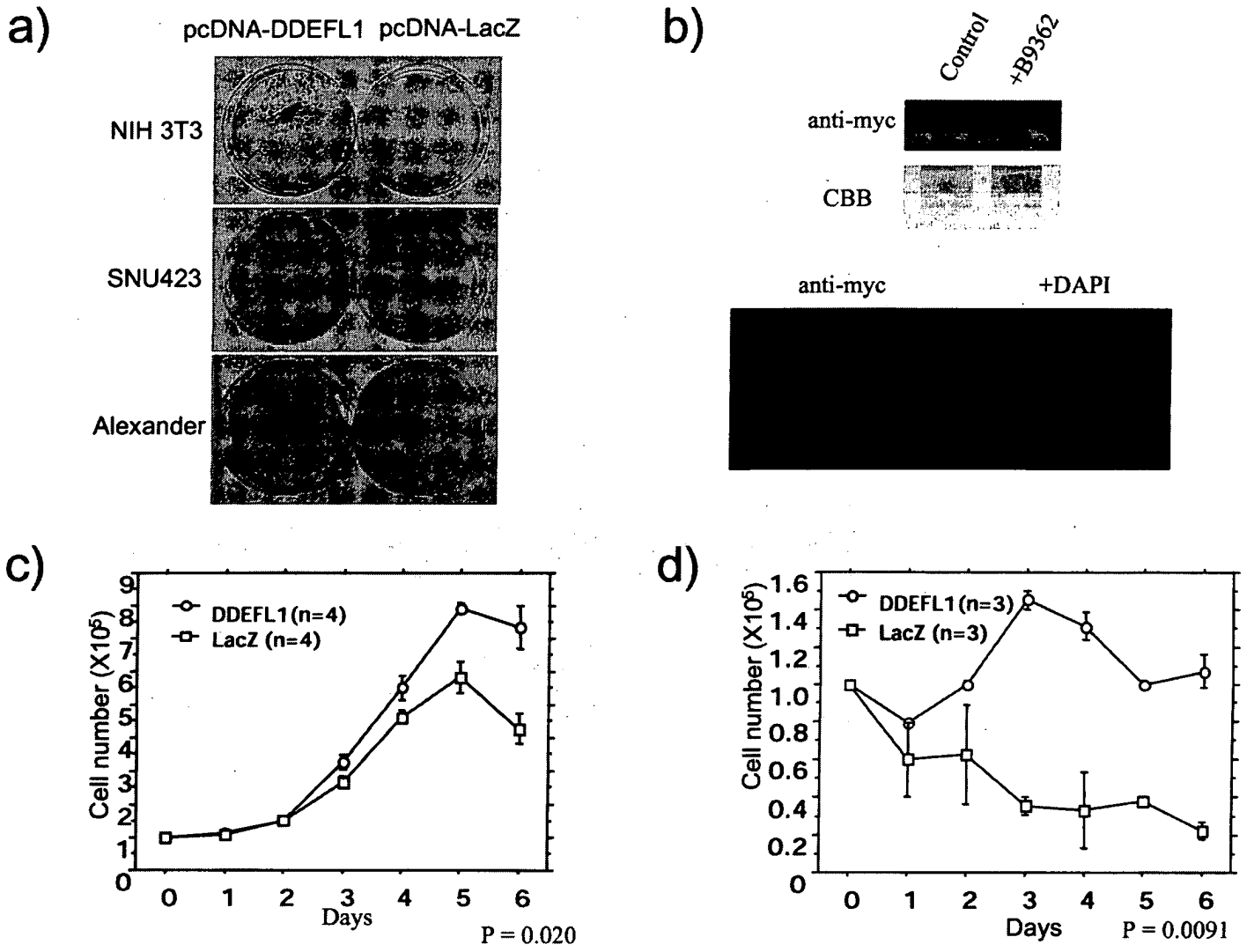

a)
pcDNA-DDEFL1
pcDNA-LacZ
NIH 3T3
SNU423
Alexander
b)
Control
+B9362
anti-myc
CBB
anti-myc
+DAPI
c)
DDEFL1 (n=4)
LacZ (n=4)
Cell number (X10⁵)
Days
P = 0.020
d)
DDEFL1 (n=3)
LacZ (n=3)
Cell number (X10⁵)
Days
P = 0.0091

Figure 5

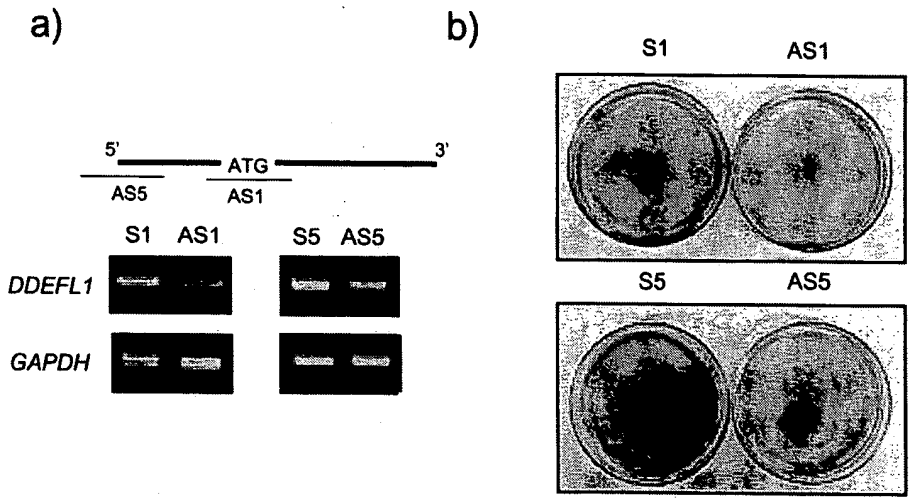
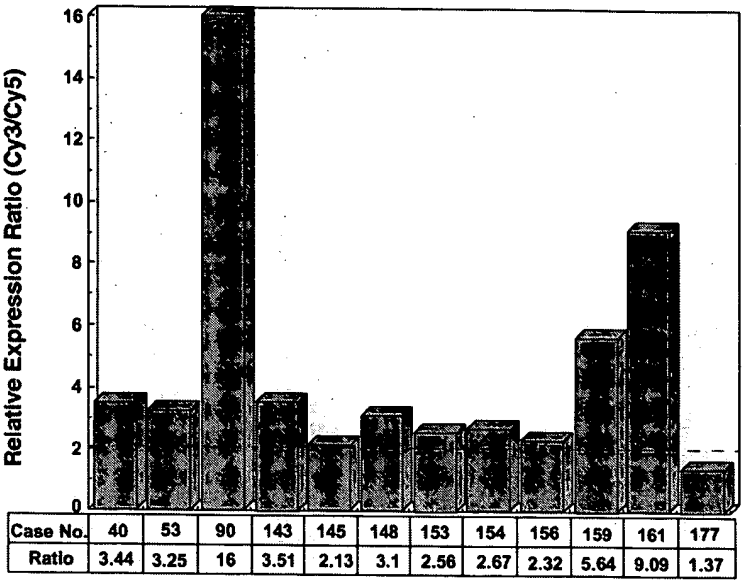


Figure 6

a



Case No.	40	53	90	143	145	148	153	154	156	159	161	177
Ratio	3.44	3.25	16	3.51	2.13	3.1	2.56	2.67	2.32	5.64	9.09	1.37

b

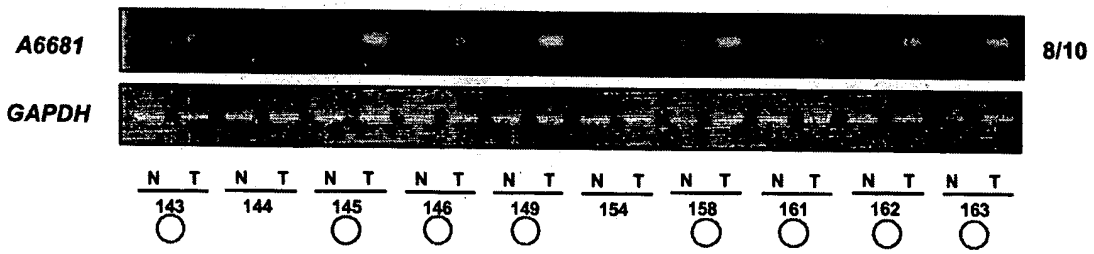


Figure 7

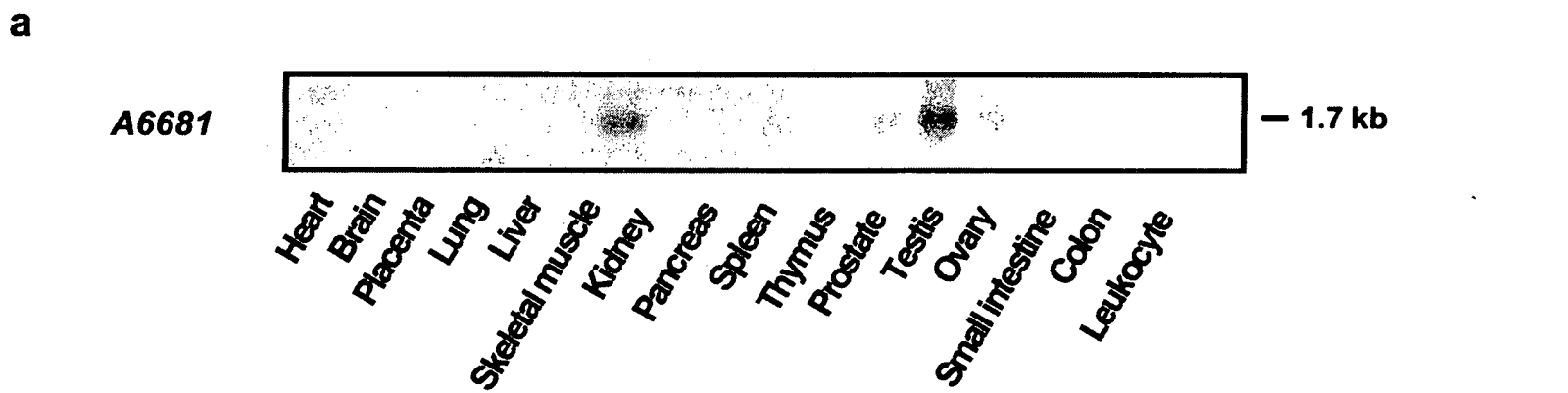

a
A6681
1.7 kb
Heart
Brain
Placenta
Lung
Liver
Skeletal muscle
Kidney
Pancreas
Spleen
Thymus
Prostate
Testis
Ovary
Small intestine
Colon
Leukocyte
b
1
50
100
150
200
250
300
350
400
428
zf-MYND
SET

Figure 8

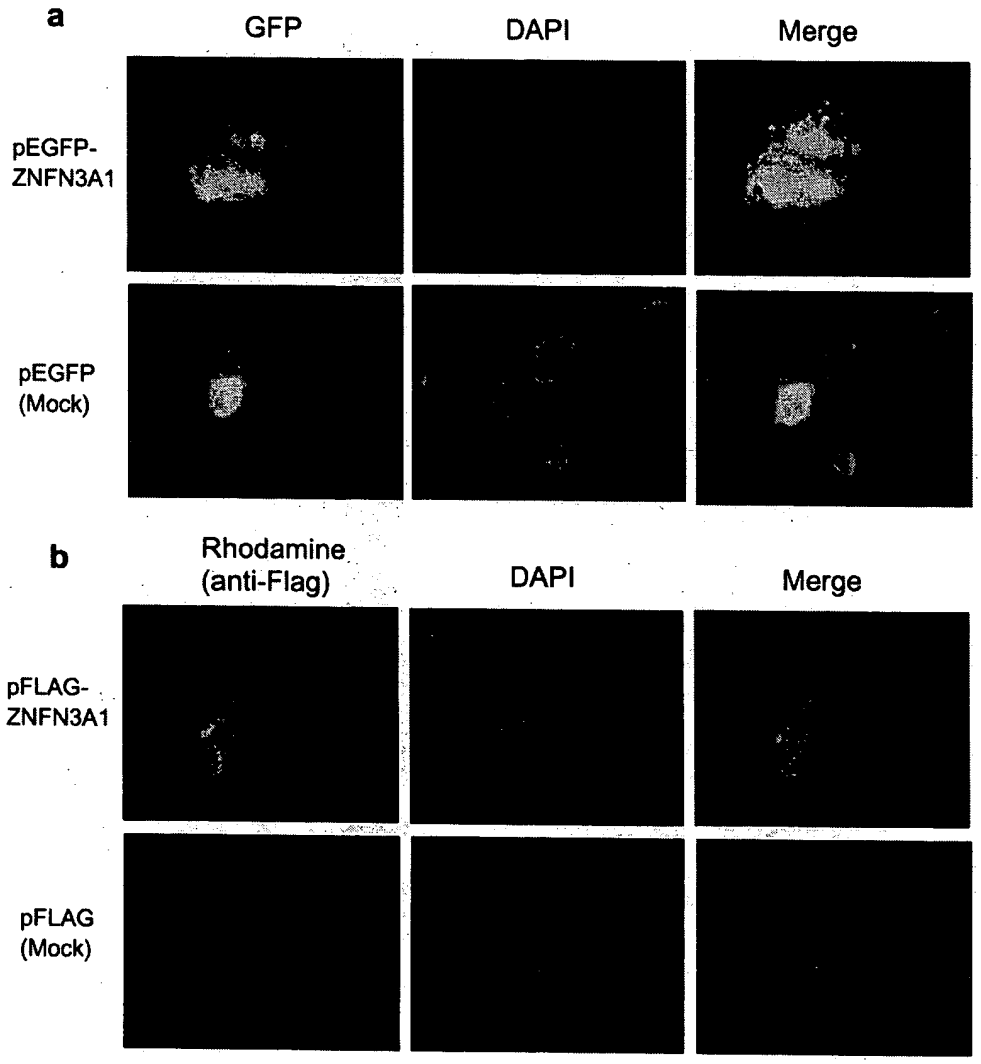

a
GFP
DAPI
Merge
pEGFP-ZNFN3A1
pEGFP (Mock)
b
Rhodamine (anti-Flag)
DAPI
Merge
pFLAG-ZNFN3A1
pFLAG (Mock)

Figure 9

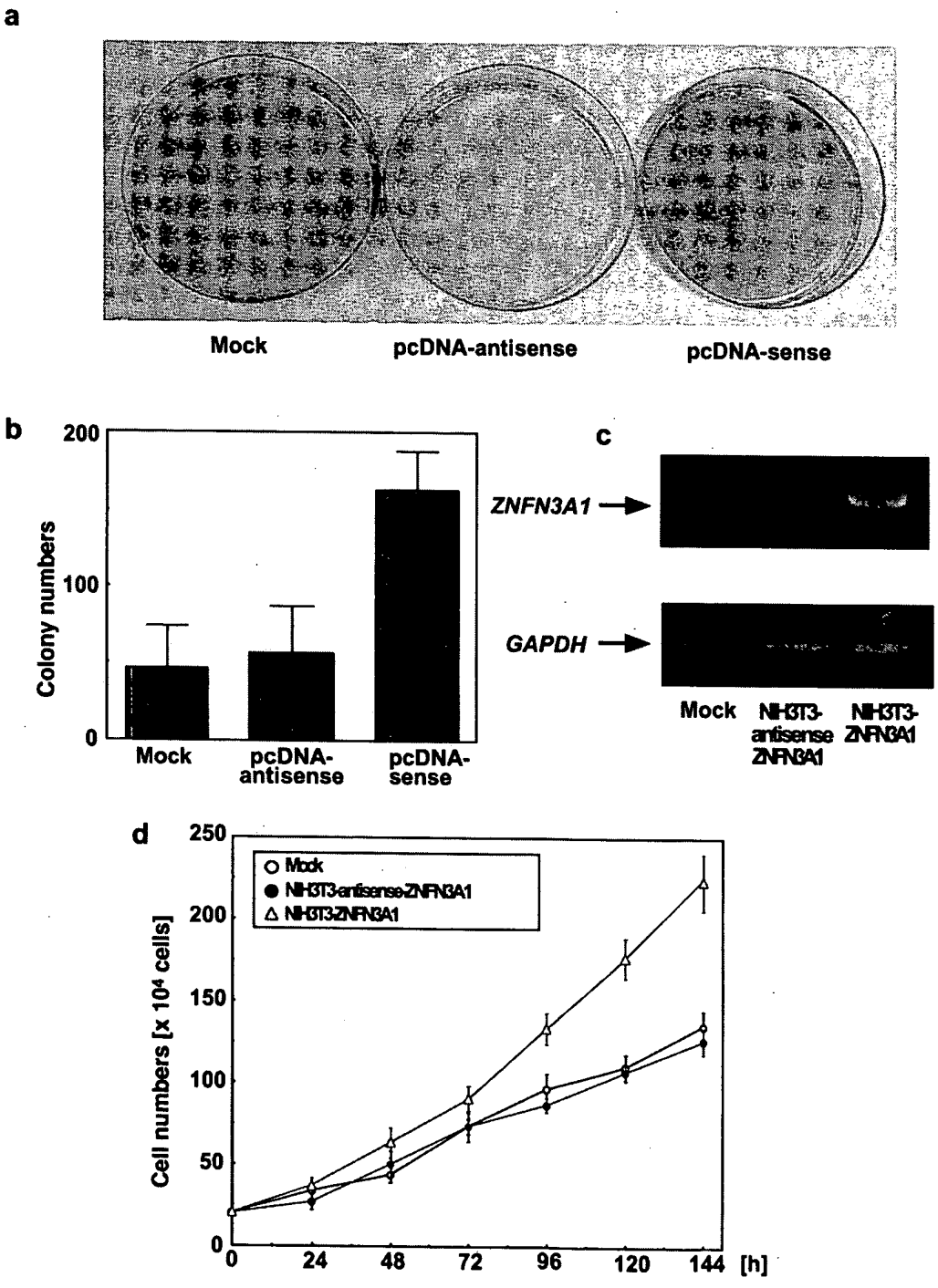

a
Mock
pcDNA-antisense
pcDNA-sense
b
Colony numbers
200
100
0
Mock
pcDNA-antisense
pcDNA-sense
c
ZNFN3A1
GAPDH
Mock
NIH3T3-antisense ZNFN3A1
NIH3T3-ZNFN3A1
d
Cell numbers [x 10⁴ cells]
250
200
150
100
50
0
Mock
NIH3T3-antisense-ZNFN3A1
NIH3T3-ZNFN3A1
0
24
48
72
96
120
144
[h]

Figure 10

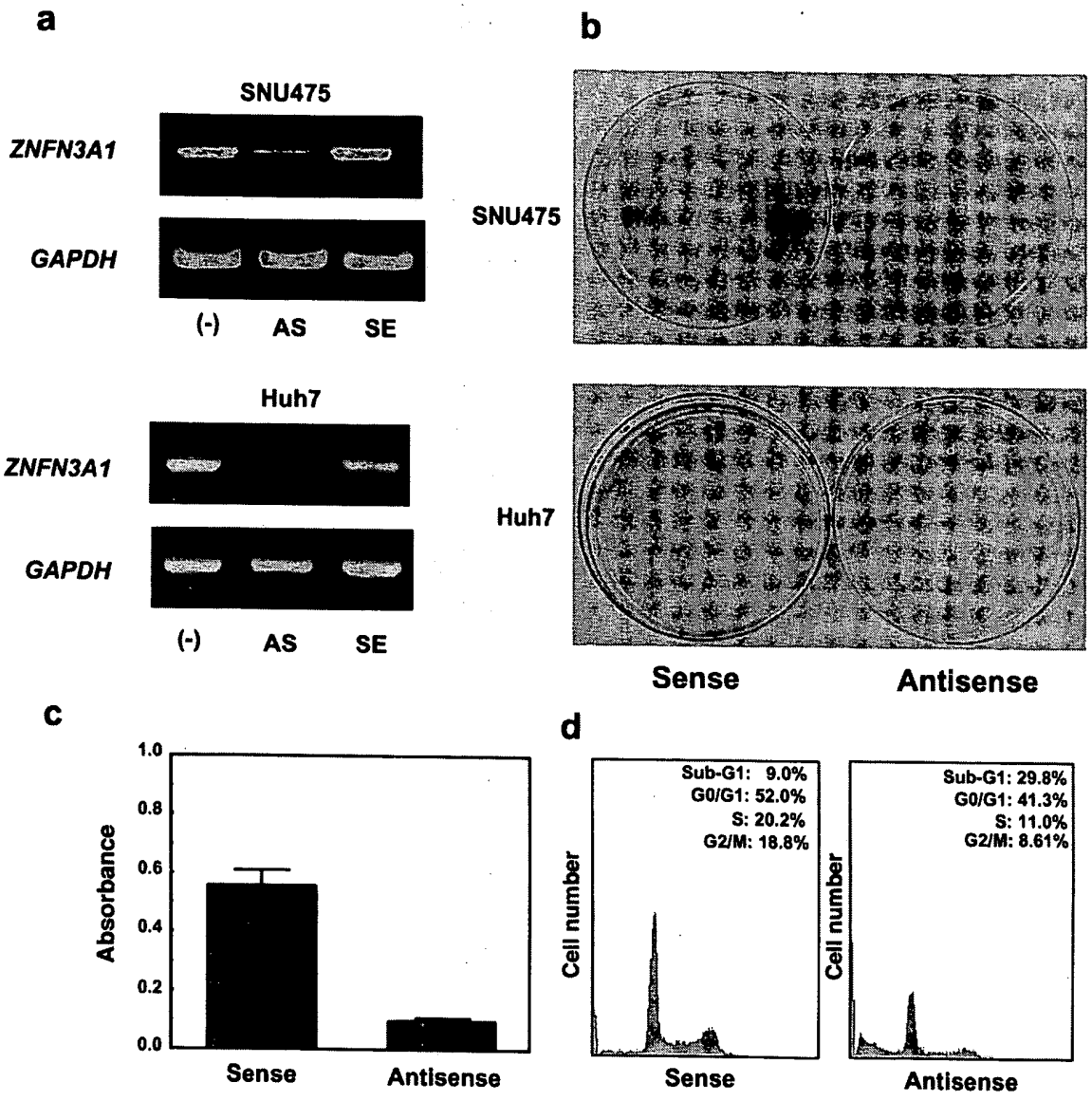
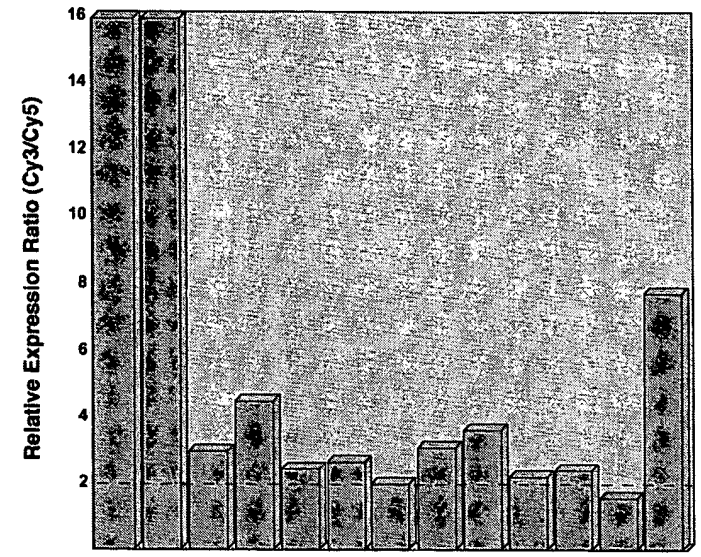

a
SNU475
ZNFN3A1
GAPDH
(-) AS SE
Huh7
ZNFN3A1
GAPDH
(-) AS SE
b
SNU475
Huh7
Sense
Antisense
c
Absorbance
1.0
0.8
0.6
0.4
0.2
0.0
Sense
Antisense
d
Sub-G1: 9.0%
G0/G1: 52.0%
S: 20.2%
G2/M: 18.8%
Cell number
Sense
Sub-G1: 29.8%
G0/G1: 41.3%
S: 11.0%
G2/M: 8.61%
Cell number
Antisense

Figure 11.

a



Case No.	53	90	104	123	143	145	146	148	153	154	161	174	177
Ratio	16.00	16.00	2.991	4.481	2.460	2.656	1.970	3.098	3.631	2.208	2.387	1.551	7.733

b

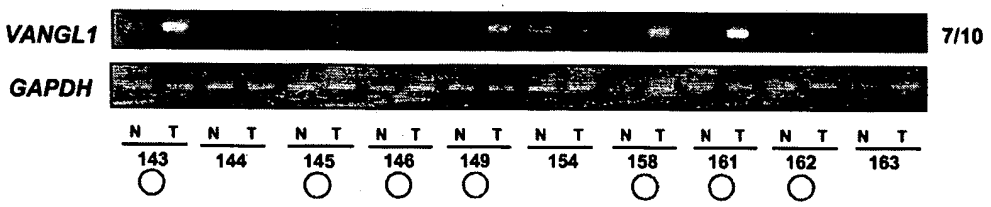
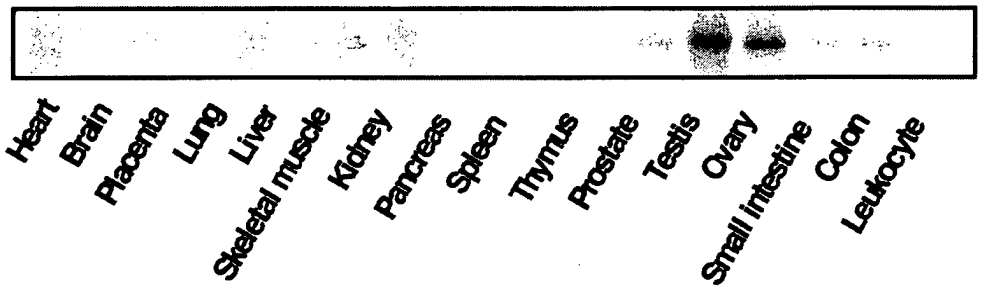


Figure 12.

a


Heart
Brain
Placenta
Lung
Liver
Skeletal muscle
Kidney
Pancreas
Spleen
Thymus
Prostate
Testis
Ovary
Small intestine
Colon
Leukocyte

b

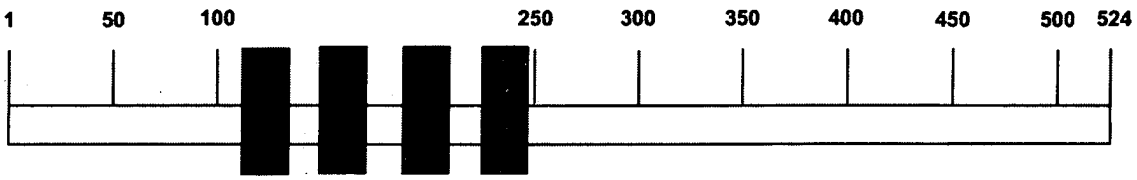

1
50
100
250
300
350
400
450
500
524

Figure 13.

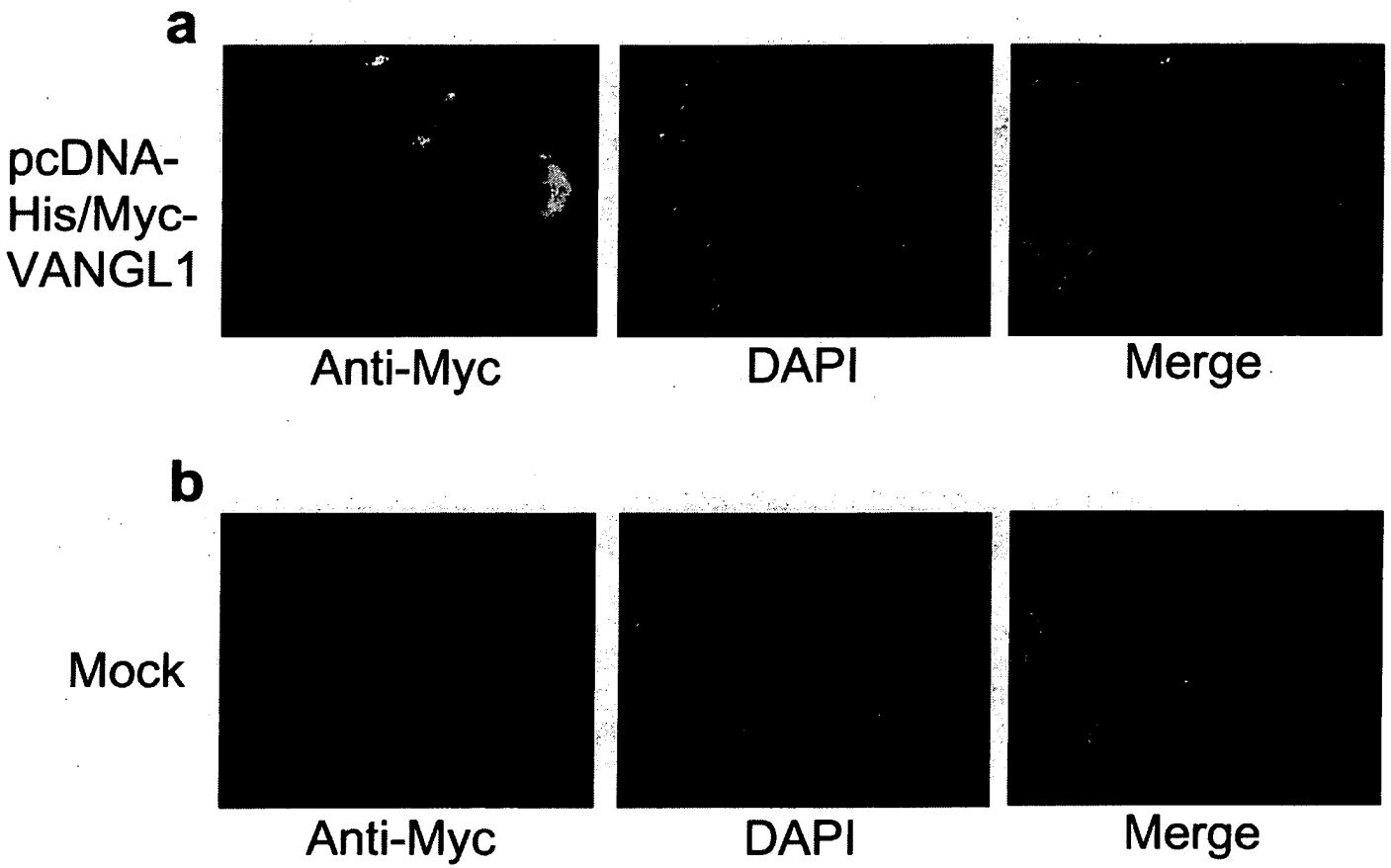

a
pcDNA-His/Myc-VANGL1
Anti-Myc
DAPI
Merge
b
Mock
Anti-Myc
DAPI
Merge

Figure 14.

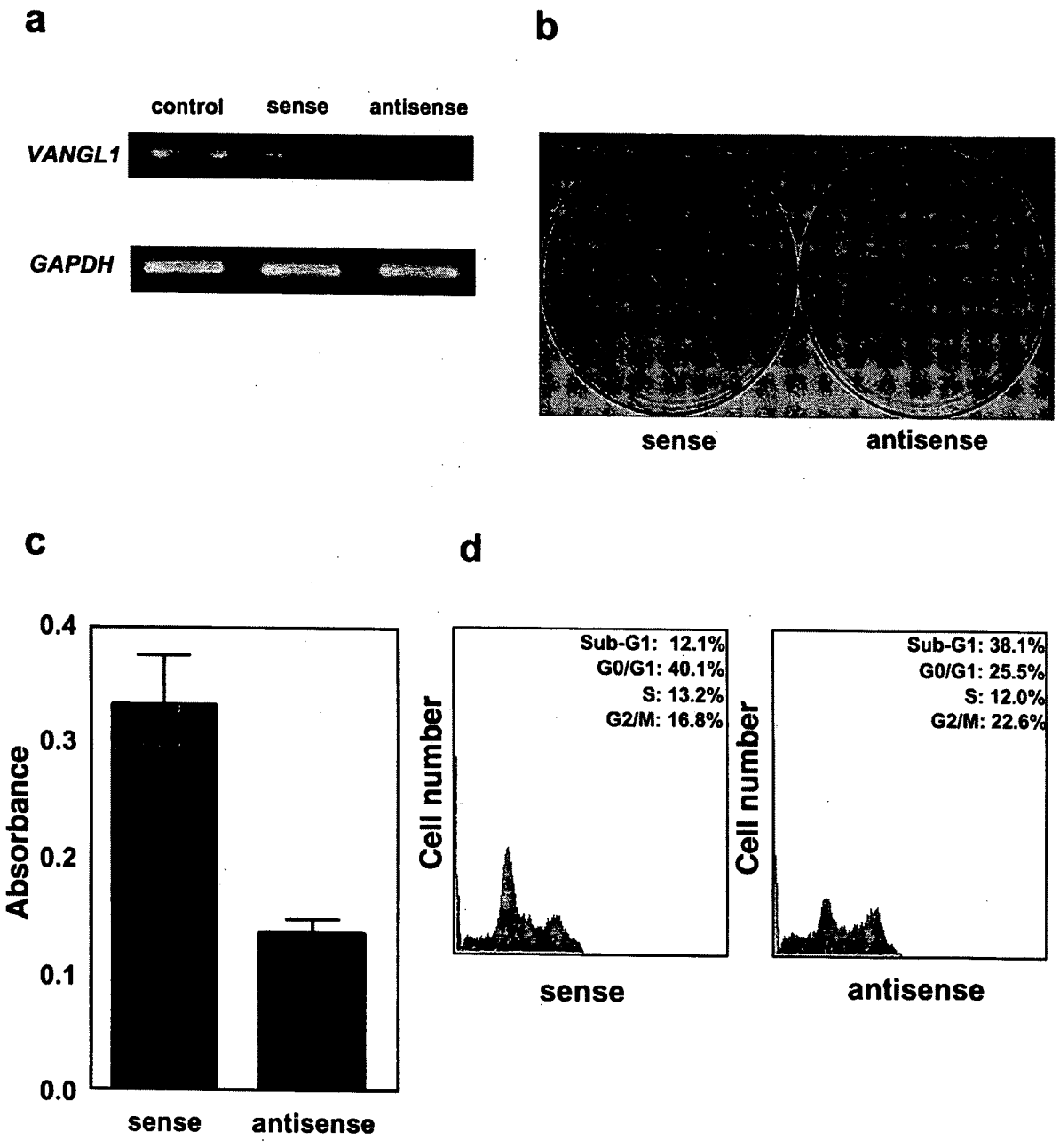
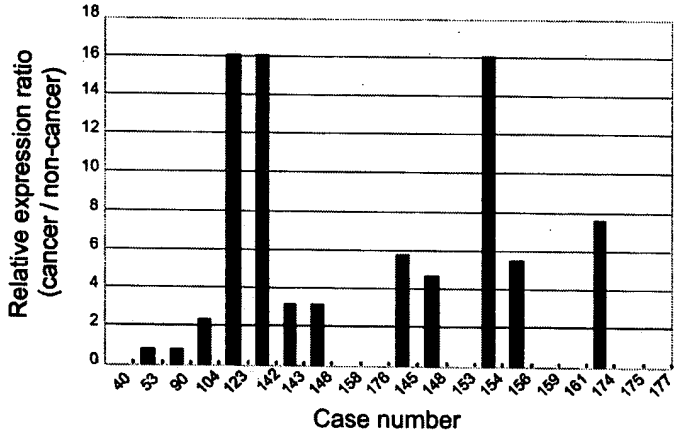
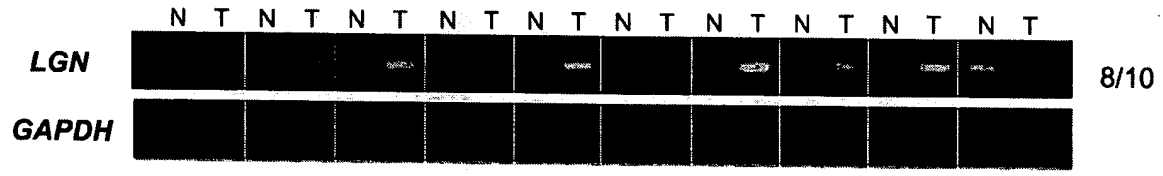

a
control
sense
antisense
VANGL1
GAPDH
b
sense
antisense
c
Absorbance
0.4
0.3
0.2
0.1
0.0
sense
antisense
d
Cell number
Sub-G1: 12.1%
G0/G1: 40.1%
S: 13.2%
G2/M: 16.8%
sense
Cell number
Sub-G1: 38.1%
G0/G1: 25.5%
S: 12.0%
G2/M: 22.6%
antisense

Figure 15

a



b



8/10

Figure 16

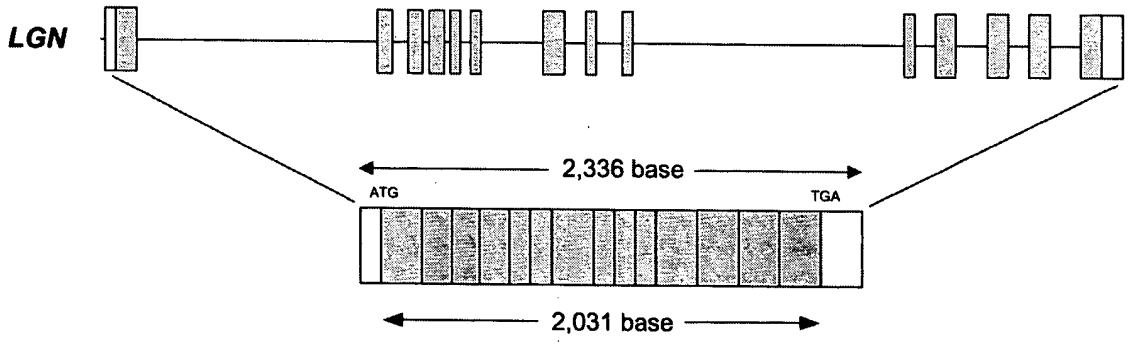
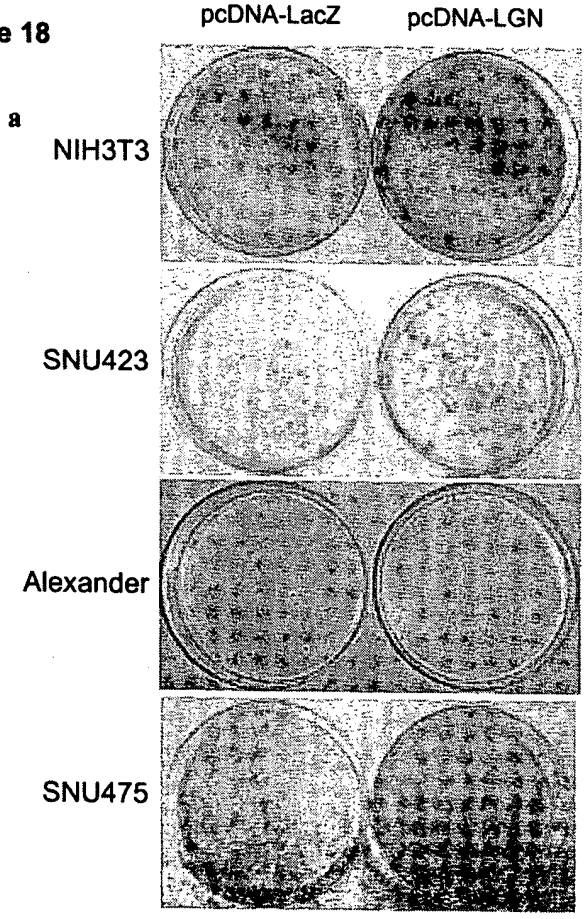


Figure 17

Figure 18


pcDNA-LacZ
pcDNA-LGN
a
NIH3T3
SNU423
Alexander
SNU475

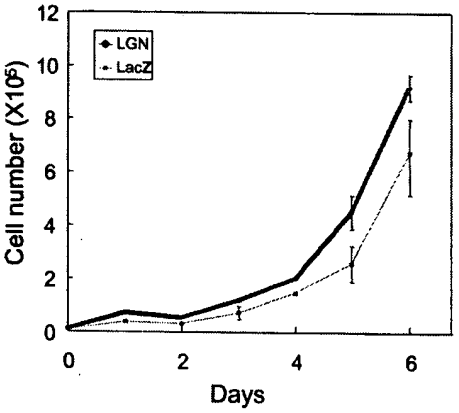

b
LGN
LacZ
Cell number (X10⁵)
0
2
4
6
8
10
12
0
2
4
6
Days

Figure 19